Development of Phase 2 Industry Codes under the Online Safety Act

eSafety Position Paper

July 2024

Australian Government | eSafetyCommissioner

eSafety.gov.au

Acknowledgement

eSafety acknowledges the Traditional Custodians of Country throughout Australia and their continuing connection to land, waters and community. We pay our respects to Aboriginal and Torres Strait Islander cultures, and to Elders past, present and emerging. Given the global nature of the internet, eSafety also acknowledges the inherent and continuing rights of indigenous people across the globe.

Content warning

This report frequently refers to pornography and sexually explicit content. It also references:

- crime and violence
- drug use and drug-related content
- suicide and self-harm
- eating disorders
- sexual violence.

1800 Respect: 1800 737 732
Qlife: 1800 184 527
Lifeline: 13 11 14

Contents

1. Executive summary

Overview and mandate

The eSafety Commissioner (**eSafety**) is Australia's independent online safety regulator. With a mandate to minimise online harms, eSafety aims to ensure safer, more positive online experiences for all Australians. eSafety exercises its legislative powers to prevent access and exposure to illegal and harmful online content and activities. It also investigates and remediates harm through legislated complaints schemes.

Legislative framework

Under the ***Online Safety Act*** *2021* (Cth) (the **OSA**), which commenced on 23 January 2022, the Commissioner can request industry bodies to develop codes to regulate illegal and restricted online material across eight sections of the online industry. If these codes meet the statutory requirements, the Commissioner can register them, making them binding on all industry participants. If a code fails to meet these requirements, eSafety can develop an enforceable industry standard for that section of the online industry instead, to ensure appropriate protections are in place for the community.

Development phases

The development of these important codes and standards has occurred in two phases.

In September 2021, eSafety published an initial position paper (**September 2021 Position Paper**) to guide the industry in developing the first phase of codes (**Phase 1 Codes**). These codes apply to 'class 1' material, such as child sexual exploitation and pro-terror content. In April 2022, eSafety issued notices to six industry bodies requesting they develop the Phase 1 Codes. The industry-developed Phase 1 Codes for five industry sections were registered in June 2023 and came into effect in December 2023. A sixth industry code was registered in September 2023 and came into effect in March 2024. Following the Commissioner's decision not to register the remaining two codes, eSafety developed standards for those industry sections, which were registered in June 2024 and will take effect in December 2024. For more details about Phase 1, see chapter 4.

Since November 2023, eSafety has been engaging with industry on the proposed approach to the second phase of industry codes (**Phase 2 Codes**). These codes aim to protect Australian children from access or exposure to online pornography and other 'class 2' material, which is identified under the National Classification Scheme as being inappropriate for children. More broadly, the Phase 2 Codes should ensure all Australian end-users have effective tools and options to limit their exposure to class 2 material if they choose not to engage with it.

Phase 2 Codes and development

Now that Phase 1 is finalised, eSafety has issued section 141 Notices to five of the six industry bodies (**Notice Recipients**) involved in drafting the Phase 1 Codes, requesting they begin drafting the Phase 2 Codes. For details, see chapter 5.

This position paper aims to support this process by outlining eSafety's expectations for developing Phase 2 Codes. These expectations are informed by lessons learned from Phase 1 and eight months of preliminary Phase 2 engagement with representative industry bodies and individual industry participants. It sets out eSafety's positions on the matters to be addressed (chapter 8), the measures which may be included (chapters 8 and 9), and timing (chapters 5 and 10).

Content categorisation and approach

The position paper also explains eSafety's approach to the different types of content which fall within the definition of class 2 material (chapter 7). It highlights content we believe warrants stronger measures in Phase 2 Codes and content likely to be more context-dependent, for which scalable measures may be more challenging.

Principles and compliance measures

Throughout this position paper, eSafety sets out expectations and suggestions about principles and compliance measures to guide industry's development of Phase 2 Codes. This includes appropriate community safeguards. The measures outlined in this position paper highlight the importance of implementing age assurance, as it is essential for industry participants to recognise when an end-user is a child to activate appropriate protective measures. eSafety also presents case studies in Annexure A to show how these measures could be adopted.

The drafting of codes in the first instance is a matter for industry. The responsibility for drafting codes includes conducting stakeholder consultation and providing eSafety with progressive draft codes in accordance with indicative targets in the Notices issued to industry bodies.

Timing

eSafety expects industry to work with urgency to submit codes for consideration before the end of 2024. eSafety considers this timeframe is reasonable given the community interests and benefits in having appropriate safeguards in place as soon as possible, together with the ability to leverage and build on the significant body of work and the processes that have already been established in Phase 1. We expand on this in chapters 5 and 10.

Regulatory environment for industry and collaboration

eSafety acknowledges the dynamic regulatory environment in Australia and around the world. This position paper identifies intersecting processes, such as the review of the Australian classification scheme, the OSA, and the *Privacy Act 1988* (Cth) (**Privacy Act**); the age assurance trial being run by the Department of Infrastructure, Transport, Regional Development, Communications and the Arts (**DITRDCA**); developments in international jurisdictions such as Ireland and the United Kingdom (**UK**); and the roll out of Digital IDs in Australia and across the world for more general identification online beyond age assurance. eSafety aims to align with these relevant processes wherever possible.

Industry engagement

Any Codes or Standards registered will be binding on industry participants regardless of their participation in the development process. The co-regulatory approach gives providers a unique opportunity to help shape these regulations.

One of the focuses of Phase 2 is preventing children's access to online pornography. While eSafety is not aware of any online pornography providers being part of the direct membership of the Notice Recipients, it strongly encourages them to engage with Australia's co-regulatory process and contribute to developing the Phase 2 Codes.

Underlying principles

eSafety's underlying principles for this position paper, and for Phase 2 in general, are to encourage industry to:

- **Protect and support children**: Prevent children from accessing or being exposed to harmful material online, and provide them with support to minimise harm.
- **Empower all end-users:** Provide better tools and more control over what end-users see, hear, read and listen to online.
- **Promote safety, privacy and human rights:** Enhance online safety, preserve privacy, and uphold the human rights of Australians, with the best interests of children as the paramount consideration.
- **Achieve comprehensive and effective safety:** Take steps across every section of the online industry to achieve these ends.

2. Purpose of this position paper and relationship to Phase 1

Purpose of this paper

The purpose of this position paper is to inform and guide industry's development of Phase 2 Codes. eSafety's intentions are to:

- **Provide clear guidance:** Meet industry's request for clear guidance on eSafety's expectations for Phase 2 Codes.
- **Support drafting:** Help industry draft Phase 2 Codes that provide appropriate community safeguards for matters of substantial relevance to the community.
- **Minimise unintended consequences:** Reduce risks related to privacy and access to information.
- **Minimise regulatory burden**: Minimise regulatory burdens for industry by promoting alignment with other Australian and international regulatory requirements.
- **Build on previous work:** Leverage the significant work conducted during the first phase of the codes development process to expedite Phase 2 Codes.

Similar to Phase 1, this paper proposes that the Phase 2 Codes should be supported by clear compliance measures which apply to industry participants whose services or devices are likely to be used by children to access class 2 material.

Phase 1 outcomes and example measures

The September 2021 position paper stated the desired outcomes of Phase 2 of the Codes, providing industry with advanced notice of eSafety's expectations for Phase 2. For ease of reference, we extract below the relevant outcomes for Phase 2 as set out in the September 2021 Position Paper:[1]

[1] eSafety Commissioner, *Development of Industry Codes: Position Paper*, p. 69.

Outcome 3: Industry participants proactively: • prevent access or exposure to, and distribution of, or • prevent children from accessing, or being exposed to class 1 - 1C and class 2 material
Industry participants will have scalable and effective policies, procedures, systems and technologies in place to proactively: • prevent access or exposure to, and distribution of, or • prevent children from accessing, or being exposed to, class 1 - 1C and class 2 material. **Examples of measures through which this outcome could be implemented** (depending on service or device provided and associated risk profile): - Implementation of age verification or age assurance mechanisms - Safety settings such as safe search mode are turned on by default - Internal policies and procedures that include safety risk and impact assessments, and safety review processes that specifically consider users aged under 18

Extract from September 2021 Position Paper, pg 69.

Overview of positions

This position paper is informed by lessons learned from developing Phase 1 Codes and by preliminary Phase 2 Codes discussions with industry between November 2023 and June 2024. It builds on and adjusts the 11 positions set out in the September 2021 Position Paper in the following ways:

September 2021 position paper	This position paper
Position 1: The codes will address the issues of access, exposure and distribution that are related to class 1 and class 2 material.	The Phase 2 Codes will address the two matters set out in the Notices issued on 1 July 2024, as explained in chapter 5. eSafety's suggestions in relation to the scope of content to be covered in Phase 2 Codes – building on the September 2021 position paper and incorporating subsequent feedback from industry – can be found in chapter 7.
Position 2: The application of the codes will not be limited to services provided from Australia.	This position remains the same. Consistent with the intent of the OSA, the codes will apply to online services so far as those services are provided to end-users in Australia.
Position 3: Industry associations will develop a set of common drafting principles to inform codes development.	eSafety considers industry can largely rely on the head terms developed for Phase 1, with a few adjustments as noted throughout this paper. The Phase 2 Codes should reflect the guiding principles of: • this position paper (at page 6) • the September 2021 position paper (at pages 45-46)

	• the Age Verification Background Report (at pages 30-31) • the National Classification Code and the Classification Guidelines (as reflected in the Phase 1 head terms).
Position 4: The codes will adopt an outcomes- and risk-based regulatory approach, supported by clear compliance measures which apply to industry participants whose services or devices present the greatest risk in respect of class 1 and class 2 material.	This position remains the same, with further guidance for Phase 2 Codes provided in chapter 8.
Position 5: Industry associations will prepare all codes for registration by July 2022 or adopt a phased approach to codes development. Under the phased approach, codes dealing with the most harmful content must be lodged for registration by July 2022, and codes dealing with content which is inappropriate for children must be lodged for registration by December 2022.	As set out in chapter 4, these timelines were not met. In Phase 1, industry submitted codes for registration in November 2022, 15 months after the September 2021 position paper was released. eSafety believes the equivalent process in Phase 2 can be condensed into six months, given industry can build upon the processes, head terms,[2] Codes, and Standards developed in Phase 1. Further details can be found in chapter 5.
Position 6: Industry associations will limit the number of codes developed.	In Phase 1, eSafety ultimately accepted separate codes drafted for each individual industry section where they provided appropriate community safeguards. The codes are contained in one consolidated document with shared head terms. Given the different focus of Phase 2, eSafety considers industry should revisit the possibility of creating combined codes for some industry sections. Certain codes could be combined where participants in those sections offer similar functionalities and purposes, and therefore where adopted measures are likely to be similar (see further discussion in chapter 4). SMS, RES and DIS might be usefully combined to avoid any confusion by industry participants in relation to the code that applies to them (and noting other schemes under the OSA apply equally to SMS, RES and DIS). If the Notice Recipients choose to draft separate Codes again, they should maintain the approach of consolidating them in one document with shared head terms. Drafters should also cross-check the draft codes and aim to avoid duplication, gaps in safeguards and inconsistent approaches across similar services.

[2] Consolidated Industry Codes of Practice for the Online Industry (Class 1A and Class 1B Material) - Head Terms, 12 September 2023.

Position 7: Industry associations will engage widely with participants within their industry section(s) to ensure they adequately represent each section covered by a code. and **Position 8:** Industry associations will conduct meaningful industry and public consultation.	This position remains the same, and eSafety retains the view that representation does not necessarily require membership in an industry body. eSafety in particular notes that the Notice Recipients should make efforts to ensure adequate consultation with online pornography providers for Phase 2 and encourages those providers in turn to actively engage with the mandatory consultation processes around any draft codes. We discuss this expectation in more length in chapter 10. eSafety has also received feedback from industry participants that the 30-day consultation timeframe contained in the OSA was inadequate in Phase 1. eSafety suggests industry should consult for a longer period than the legislated minimum if it would assist the code production process. Industry may need to adjust their internal timeframes for code drafting accordingly.
Position 9: Industry associations will engage with eSafety throughout the codes development process.	This position remains the same. eSafety appreciates the constructive way industry has engaged with both phases of the codes development process to date. To support this position and encourage productive and efficient consultation between eSafety and industry associations, eSafety has also included an indicative target for the provision of draft codes to eSafety by 3 October 2024 in the s 141 Notices.
Position 10: Industry participants will handle reports and complaints about class 1 and class 2 material and codes compliance in the first instance. eSafety will act as a 'safety net' if resolution of a complaint is not satisfactory.	This position remains the same. eSafety considers industry may replicate the Complaints about Code compliance provisions set out in the Phase 1 head terms.
Position 11: The codes will include a review mechanism.	This position remains the same. eSafety has encouraged industry to adopt 'supportive' measures in the Phase 2 Codes, such as reporting functions, in addition to the main protective measures for Phase 2 suggested in this position paper. eSafety considers industry may replicate the Code review provisions set out in the Phase 1 head terms.

3. Background

Industry Codes and Standards under the Online Safety Act

Part 9, Division 7 of the OSA provides for developing industry codes and standards regulating the online activity of eight sections of the online industry.[3] These are set out in section 135:

- **App distribution services** - APP - services distributing apps that can be downloaded and accessed by end-users in Australia. For example, app stores/marketplaces.[4]
- **Equipment services** - EQP - manufacturers, suppliers, maintainers and installers of equipment that is used to access online services[5] such as mobile phones; laptops; tablets; internet-enabled devices (such as smart TVs and gaming consoles); immersive technologies (such as virtual reality headsets); wi-fi routers. This section of the online industry includes manufacturers of these devices, operating system providers, as well as businesses and retail outlets that install, sell and/or repair/maintain such devices.
- **Hosting services** - HOS - services which host stored material in Australia (for example, a service with data centres located in Australia).[6]
- **Internet carriage services** - ICS - a listed carriage service that enables end-users to access the internet.[7]
- **Internet search engine services** - SES -electronic services designed to collect, organise (index) and/or rank information on the world wide web (WWW) in response to end-user queries and return search results to end-user queries.[8]
- **Relevant electronic services** - RES - services that can be accessed by end-users in Australia, including but not limited to instant messaging services; Short

[3] In this position paper, where eSafety refers to 'online services' it can be taken as a reference to an online activity within the meaning of s 134 of the OSA by a member of a section of online industry.
[4] *Online Safety Act 2021* (Cth), s 5 (definition of 'App Distribution Services'). Excludes links to an app and download of apps from third party websites.
[5] 'Online services' are social media services, relevant electronic services, designated internet services and internet carriage services: *Online Safety Act 2021* (Cth), s 135(2)(h).
[6] *Online Safety Act 2021* (Cth), s 17. eSafety notes that 'internet service provider' is defined in s 19 of the OSA as 'a person who supplies, or proposes to supply, an internet carriage service to the public.'
[7] *Online Safety Act 2021* (Cth), s 5.
[8] Excludes search functionality within platforms where content or information can only be surfaced from that which has been generated/uploaded/created within the platform itself and not from the WWW more broadly.

Message Services and Multimedia Message Services; chat services; online multi-player gaming services; email services; online dating services; enterprise messaging services.[9]

- **Social media services** – 

- services enabling online social interaction between end-users. For example, social networks; public media sharing networks; discussion forums; consumer review networks.[10]
- **Designated internet services** - 

- services allowing end-users in Australia to access material using an internet carriage service, where the service is not a Social Media Service or Relevant Electronic Service. For example, file storage services managed by end-users in Australia, websites and apps.[11]

References to the '**technology stack'** in this position paper pertain to the eight sections of the online industry set out in section 135 of the OSA and described above. The technology stack differs from technological '**ecosystems**' which refer to the interconnections between online service providers' products and services which may affect how end-users interact with the technology stack. Ecosystems are discussed in further depth in chapter 8.

Section 141 of the OSA allows for the Commissioner to issue notices to industry bodies or associations representing one or more sections of the online industry, requesting that they develop industry codes that deal with a wide range of matters. Section 138 of the OSA sets out a non-exhaustive list of example matters that industry codes may address.[12]

Section 145 allows the Commissioner to determine an industry standard which applies to participants in a particular section of the online industry if certain conditions related to the request under section 141 are not met. These conditions include:

- **Non-compliance with the request:** The request from the Commissioner is not complied with.
- **Inadequate safeguards:** A draft code developed pursuant to the request does not contain appropriate community safeguards to deal with one or more matters specified in the request.

---

[9] *Online Safety Act 2021* (Cth), s 13A.
[10] *Online Safety Act 2021* (Cth), s 13. Online social interaction does not include online business interaction.
[11] *Online Safety Act 2021* (Cth), s 14.
[12] Examples of matters include but are not limited to, providing end-users with technological solutions and advice to help them limit their and their children's access to class 1 and 2 material: see *Online Safety Act 2021* (Cth), s 138(3)(f), s 138(3)(i)—(j).

- **Missed targets:** Any indicative targets specified in the notice for achieving progress in developing the code were not met.
- **Refusal of registration:** The request is complied with, but the Commissioner subsequently refuses to register the Code.

Once the codes or standards are registered, compliance is mandatory. eSafety has a range of compliance and enforcement powers it can apply where an industry participant does not comply with the codes or standards. In line with its *Regulatory Posture and Regulatory Priorities* and *Compliance and Enforcement Policy*, eSafety performs its regulatory functions and exercises its powers in a fair, transparent and proportionate way to promote compliance, prevent and remediate online harm, and drive continuous improvement in safety by online service providers.[13]

At the time of writing this position paper, an independent statutory review of the OSA (**the Review**) is being conducted. The Review is broad-ranging and will examine and report on the effectiveness of the OSA in achieving its overarching objectives. This includes whether amendments should be made to the legislative framework governing industry codes.[14] More information about the Review and how it relates to the development of Phase 2 Industry Codes is in chapter 6 of this paper.

Classes of material under the Online Content Scheme

eSafety's mandate to register industry codes or standards falls under Part 9 of the OSA, which sets out the Online Content Scheme applying to class 1 and class 2 material.

The Online Content Scheme defines class 1 and class 2 material by referring to the National Classification Scheme (**NCS**). The NCS is a cooperative arrangement between the Australian Government and state and territory governments. Its purpose is to classify films, publications and computer games.

The NCS is implemented through the *Classification (Publications, Films and Computer Games) Act 1995* (Cth) (**Classification Act**) and complementary state and territory enforcement legislation. The Classification Act enlivens the *National Classification Code (May 2005)* and several sets of guidelines (collectively known as the **Classification Guidelines**). These guidelines determine how films, publications and computer games should be classified. Specifically, these are the *Guidelines for the Classification of Films*

[13] eSafety Commissioner, *Compliance and Enforcement Policy*, December 2021, pg. 6-8. See also *Phase 1 Industry Codes (Class 1A and Class 1B Material) Regulatory Guidance*, pg. 36-42.
[14] Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Terms of Reference - Statutory Review of the Online Safety Act 2021*, February 2024.

2012 (**Film Guidelines**), the *Guidelines for the Classification of Computer Games 2012* (**Computer Games Guidelines**), and the *Guidelines for the Classification of Publications 2005* (**Publications Guidelines**).

The NCS is currently undergoing a two-staged review process. Stage 1 reforms will begin in 2024, while public consultations on stage 2 reforms concluded in May 2024. This is examined in more detail in chapter 6. The Online Content Scheme applies to materials which would be rated R18+ (or Category 1 in the Publications Guideline) or higher.

Classification of Content

Online Safety Act	Content	Classification Act/National Classification Code
Class 1	Film Publication Computer game Any other material*	Refused Classification (RC)
Class 2	Film Any other material (excluding computer games)*	X18+
	Publication	Category 2 restricted
	Film Computer game Any other material *	R18+
	Publication	Category 1 restricted
* Under the Online Safety Act, material that is not a film, computer game or publication is to be classified in a corresponding way to the way in which a film would be classified		

Extract from September 2021 Position Paper, pg 19.

4. Phase 1 Codes and Standards

Classes of material and phased approach

In the September 2021 position paper, eSafety addressed the context in which the National Classification Code and Guidelines were created. Recognising the differences in the modern online environment,[15] eSafety proposed subcategories for class 1 and class 2 material and encouraged industry bodies to adopt a two-phased approach when developing industry codes:[16]

Phase 1: Focuses on high-end class 1 material (1A and 1B) including child sexual exploitation material and pro-terror content. The primary goal here is to prevent or restrict access to material that poses harm to people of all ages.

Phase 2: Covers online pornography (class 1C and class 2 material) and other class 2 content. This phase aims to prevent children from accessing age-inappropriate material and offers users effective tools to manage exposure to class 2 content they do not want to see.

Phase	Class Subcategory	Material	National Classification Scheme
Phase 1	Class 1A	• CSEM – Child sexual exploitation material. Material that promotes or provides instruction of paedophile activity. • Pro-terror content – Material that advocates the doing of a terrorist act (including terrorist manifestos). • Extreme crime and violence – Material that describes, depicts, expresses or otherwise deals with matters of extreme crime, cruelty or violence (including sexual violence) without justification.[17] For example, murder, suicide, torture and rape. Material that promotes, incites or instructs in matters of extreme crime or violence.	• Class 1 • Refused Classification
Phase 1	Class 1B	• Crime and violence – Material that describes, depicts, expresses or otherwise deals with matters of crime, cruelty or violence without justification. Material that promotes, incites or instructs in matters of crime or violence.	• Class 1 • Refused Classification

[15] The NCS is currently undergoing a reform process: see chapter 6 for further discussion. eSafety understands that these reforms are unlikely to affect the definitions of class 1 or class 2 material directly in the immediate future.
[16] eSafety Commissioner, *Development of Industry Codes: Position Paper*, September 2021, p. 53.
[17] Reference to 'without justification' highlights that the nature of the material must be considered, including its literary, artistic, or educational merit and whether it serves a medical, legal, social or scientific purpose. Section 11 of the *Classification (Publications, Films and Computer Games) Act 1995* (Cth) outlines matters to be taken into account in making a decision on classification.

		• Drug-related content – Material that describes, depicts, expresses or otherwise deals with matters of drug misuse or addiction without justification. Material which includes detailed instruction or promotion of proscribed drug use.	
Phase 2	Class 1C	• Online pornography – material that describes or depicts specific fetish practices or fantasies.	• Class 1 • Refused Classification
Phase 2	Class 2A	• Online pornography – other sexually explicit material that depicts actual (not simulated) sex between consenting adults.	• Class 2 • X18+
Phase 2	Class 2B	• Online pornography – material which includes realistically simulated sexual activity between adults. Material which includes high-impact[18] nudity. • Other high-impact material which includes high-impact sex, nudity, violence, drug use, language and themes. 'Themes' includes social Issues such as crime, suicide, drug and alcohol dependency, death, serious illness, family breakdown and racism. • Simulated gambling in computer games (from September 2024).[19]	• Class 2 • R18+

For ease of reference, within this paper where eSafety refers to 'class 2' material, this is intended to also encompass the 'class 1C' material detailed above. This is because eSafety has proposed the application of the same measures for online pornography material which goes across class 1C and class 2 (see discussion at **Table 2** below).

eSafety proposed this phased approach because each phase serves distinct public policy objectives.

Phase 1 prioritises combating the most severe forms of harmful material, ensuring robust measures are in place to restrict its access and distribution.

Phase 2 builds on extensive research and consultations conducted between August 2021 and March 2023 to produce eSafety's *Roadmap for age verification and complementary measures to prevent and mitigate harms to children from online pornography* (**Age Verification Roadmap**) and *Background Report*. This included engagement with stakeholders in the adult industry, who will be affected by Phase 2 Codes governing online pornography.

[18] Impact may be higher where content is detailed, accentuated, or uses special effects, prolonged, repeated frequently, realistic or encourages interactivity.

[19] Actual gambling is addressed and regulated by the Australian Communications and Media Authority under the *Interactive Gambling Act 2001* (Cth), and by state and Commonwealth governments and agencies under the National Consumer Protection Framework for Online Wagering.

Other advantages of the two-phased approach identified in the September 2021 position paper included:

- focusing efforts on preventing or mitigating the most serious material first, thereby implementing meaningful measures swiftly
- leveraging existing research and stakeholder input to inform industry actions
- enhancing public understanding of the codes by using commonly understood terminology
- remaining consistent with the legislative requirements of the OSA.

Notices to representative industry associations

In April 2022, after several months of discussion with industry, eSafety issued s 141 Notices to the following industry bodies, requesting them to develop codes addressing class 1A and class 1B material:

Industry section	Representative industry bodies associations for Phase 1 Industry Codes
App Distribution Services	• Communications Alliance Ltd (**CA**) • Digital Industry Group Inc. (**DIGI**) • Interactive Games and Entertainment Association (**IGEA**)
Social Media Services	• Communications Alliance Ltd • Digital Industry Group Inc.
Relevant Electronic Services	• Australian Mobile Telecommunications Association (**AMTA**) • BSA | The Software Alliance (**BSA**) • Digital Industry Group Inc. • Interactive Games and Entertainment Association
Equipment Services	• Australian Mobile Telecommunications Association • Communications Alliance Ltd • Consumer Electronic Suppliers Association (**CESA**) • Interactive Games and Entertainment Association
Internet Carriage Services	• Communications Alliance Ltd
Search Engine Services	• Communications Alliance Ltd • Digital Industry Group Inc.
Hosting Services	• BSA | The Software Alliance • Communications Alliance Ltd
Designated Internet Services	• Australian Mobile Telecommunications Association • BSA | The Software Alliance • Communications Alliance Ltd • Consumer Electronic Suppliers Association

These industry bodies subsequently formed a steering group dedicated to drafting Phase 1 Codes.

The section 141 Notices, issued by eSafety on 11 April 2022, directed industry representatives to take the following actions:

- Implement policies, procedures, systems and technologies to take reasonable and proactive steps to detect and prevent access or exposure to, distribution and storage of Class 1A and Class 1B material.
- Establish policies and procedures to facilitate consultation, cooperation and collaboration among industry participants to remove, disrupt or restrict Class 1A and Class 1B material.
- Ensure effective communication and cooperation with eSafety about Class 1A and Class 1B material, including complaints.
- Provide a range of technical tools and information for people to limit their access and exposure to Class 1A and Class 1B material, including for children in their care.
- Maintain clear, accessible and effective avenues for making reports and complaints about the handling of reports about Class 1A and Class 1B material (and associated user accounts), including policies, procedures and technology to deal with these reports and complaints.
- Publish accessible and plain language policies, procedures and guidelines explaining how the online service handles Class 1A and Class 1B material, including informing end-users in Australia about safety associated safety concerns.
- Publish annual reports about Class 1A and Class 1B material, and report annually on compliance with the industry codes.[20]

Process and timeline

The OSA sets out the process and minimum timeframes for certain steps in the development and registration of industry codes. The timeline table below provides a concise overview of the practical code development process for Phase 1 Industry Codes.[21]

Timeframe or date	Development milestone or event
23 June 2021	The Online Safety Act 2021 (Cth) is passed (commencing in January 2022) and establishes the Industry Codes and Standards scheme.

[20] These matters were common across notices relating to the eight industry sections. An additional matter for Hosting Services specified measures for industry participants to have policies, procedures, systems and technologies to take reasonable and proactive steps to limit the hosting to Class 1A material and Class 1B material in Australia.

[21] *Online Safety Act 2021* (Cth), s 137, 140, 145.

September 2021	eSafety publishes *Development of Industry Codes under the Online Safety Act* (Position Paper) setting out a two-phase approach of industry codes, commencing with Phase 1 of codes development for Class 1A and 1B material.
Late 2021	An industry steering group is formed, and work commences on drafting industry codes.
February 2022	First draft industry codes provided to eSafety by industry steering group for feedback.
11 April 2022	eSafety issues notices to representatives of online industry sections to formally request they develop codes and submit them to eSafety by 9 September 2022.
May-June 2022	Industry steering group provides eSafety drafts of industry codes for preliminary views and comment (not for registration).
23 June 2022	eSafety issues a second set of notices extending the date by which codes should be submitted to eSafety to 18 November 2022.
August 2022	Steering group shares with eSafety draft industry codes intended for public consultation.
September 2022	Steering group conducts public consultation for 30 days (88 submissions received).
18 November 2022	Steering group submits revised draft industry codes, incorporating public feedback, to eSafety for registration.
Early 2023	eSafety issues preliminary views that draft industry codes submitted for registration are unlikely to be registered. The steering group is invited by eSafety on 9 February to respond to these preliminary views or resubmit draft industry codes for registration by 9 March. The steering group is granted an extension until 31 March in order to conduct a second round of public consultation (25 submissions received).
March 2023	Steering group resubmits eight industry codes for registration on 31 March.
31 May 2023	The Commissioner decides to register industry codes for five industry sections: Social Media Services; App Distribution Services; Hosting Services; Internet Carriage Services; Equipment. The Commissioner reserves decision on industry code for Internet Search Engine Services due to concerns that the code does not sufficiently capture the proposed changes to search engines to incorporate generative artificial intelligence features. The Commissioner declines to register industry codes for Relevant Electronic Services and Designated Internet Services because the codes do not provide appropriate community safeguards to deal with class 1A and class 1B material. eSafety publishes a summary of decisions.
16 June 2023	Industry codes for five industry sections are registered: Social Media Services; App Distribution Services; Hosting Services; Internet Carriage Services; Equipment.

July – September 2023	Industry code for Internet Search Engine Services is revised several times and resubmitted for registration. The Commissioner decides to register industry code for Internet Search Engine Services. Industry code for Internet Search Engine Services is registered on 12 September. eSafety publishes a summary of the Commissioner's decision on 7 September. Industry standards development commences.
November 2023	eSafety meets with DIGI, CA, AMTA and CESA about Phase 2 Codes process (with continuing meetings and correspondence about Phase 2 occurring until June 2024).
16 December 2023	Industry codes for five industry sections come into force: Social Media Services; App Distribution Services; Hosting Services; Internet Carriage Services; Equipment.
12 March 2024	Industry code for Internet Search Engine services comes into effect.
21 June 2024	Phase 1 Industry Standards for RES and DIS registered.
22 December 2024	Phase 1 Industry Standards for RES and DIS come into force.

Reasons why the Commissioner did not register Phase 1 Industry Codes for RES and DIS

The Commissioner registered submitted Codes for all industry sectors except DIS and RES.

Consequently, eSafety opted to develop industry standards for these sectors. These Standards built on industry's draft codes, addressing matters where eSafety identified a need for stronger community safeguards and integrating stakeholder feedback from public consultations.

For example, in response to industry feedback on the draft definition of 'technical feasibility', alongside considerations of enforceability and stakeholder concerns, eSafety navigated differing viewpoints. Some industry participants submitted that the concept of technical feasibility alone was insufficient to encompass broader impediments a service provider might encounter in implementing a technology to detect and remove known child sexual abuse material (**CSAM**) and known pro-terror material. Conversely, others supported tightening this provision, submitting it might render the Standards ineffective if too broadly interpreted and that this may enable industry participants to evade even technically feasible measures.

In response to these diverse perspectives, eSafety determined Standards that:

- leave 'technical feasibility' undefined, maintaining its ordinary meaning under the law, while introducing additional exceptions to the detection and removal requirements to clarify circumstances where implementing a system or technology is not reasonably practicable

- ensure accountability by giving the Commissioner the ability to require reports from service providers detailing the technical feasibility and practicability of complying with the relevant requirements in the Standards. These reports must justify any claimed exceptions and outline alternative actions taken, demonstrating and why they are appropriate in the circumstances.

The **Phase 1 Industry Standards** were registered on 21 June 2024 and will come into effect on 22 December 2024.

Key lessons

The process of developing co-regulatory Phase 1 Industry Codes was a first for online services regulated under the OSA. Several crucial lessons emerged during this milestone experience, which may help the online industry to develop Phase 2 Industry Codes. These lessons, based on eSafety's own experience and informal feedback from industry, are outlined below:

- **Setting clear expectations early for community safeguards and compliance measures:** Industry feedback highlighted a desire for eSafety to clearly define expectations for compliance measures aimed at ensuring appropriate community safeguards[22]. This included specifying the types of online harms targeted and outlining appropriate or reasonable actions expected from industry, such as proactive detection and removal of known CSAM using systems and technologies.

  For Phase 2, through this position paper eSafety aims to clarify the types of material to be addressed and the types of measures that may be appropriate to do so by:

  - setting out a detailed approach to the scope of class 2 material
  - encouraging the adoption of steps which build on Phase 1 Codes and Standards and align with existing regulatory schemes
  - proposing specific measures for each section of the industry.
- **Definitions of industry sections and categorisation of online services:** Industry participants emphasised that early identification and categorisation of online services under statutory definitions, would assist in mapping obligations in the technology stack and supply chain, determining risk thresholds, and guiding regulatory measures effectively across different sectors.

[22] The Commissioner must decide whether the industry codes submitted for registration provide appropriate community safeguards: *Online Safety Act 2021* (Cth), s 145(a)(ii).

eSafety does not have the capacity under the OSA to designate which industry section an industry participant belongs to. However, for Phase 2, eSafety has tried to be clear in this position paper about the measures it recommends for each section of industry to address each of the matters. This approach acknowledges that certain measures should be implemented across the technology stack. It also recognises the convergence of functionalities across multiple industry sections, making it more practical to adopt eSafety's proposed 'ecosystem' measures.

- **Classification of material:** Industry participants suggested eSafety should establish expectations, based on the NCS, for classifying certain types of material as class 1 or class 2, especially when the material is ambiguous or subjective. To address this, eSafety has set out in this paper a detailed approach for how industry may address the full scope of class 2 material.
- **Consultation process:** Industry feedback suggests that a consultation period longer than 30 days may be beneficial.[23] This extended period allows industry associations responsible for developing industry codes to consult with and receive submissions from stakeholders outside their membership, or bodies representing other community interests. As well, because some industry stakeholders may be less inclined to consult on draft industry codes, extra steps may be needed to encourage voluntary participation in the consultation process.

  In preliminary discussions about the Phase 2 Codes, eSafety has encouraged industry to adopt a consultation period longer than the minimum required by the OSA. eSafety believes this can be achieved within the proposed 6-month drafting period through efficient collaboration between Notice Recipients and eSafety. eSafety sets out its consultation expectations in more detail in chapter 10.
- **Structure of Phase 1 Industry Codes:** The Phase 1 Industry Codes were structured with common head terms and individual schedules for each industry section. This approach provided consistency across all industry sections, while allowing tailored approaches for each one.

  For Phase 2, eSafety believes that any industry code should apply to at least one whole industry section as identified in the OSA. The OSA does not provide for the creation of industry codes that apply only to a sub-section of an identified industry section (such as, a code only for online pornography providers). However, this does not preclude combining industry codes (which apply to at least one whole industry section), recognising the shared measures likely to be adopted by those sections.

[23] The *Online Safety Act 2021 (Cth)* requires a minimum consultation period of 30 days: s 140(3).

eSafety suggests that within an individual or joint code it may be open to industry to define specific categories of services that provide particular types of material. If industry believes that measures within a code should be applied to a specific type of online industry participant, these defined categories could then be subject to different measures or cause a service to be allocated a certain risk rating by default. This approach aligns with the Phase 1 Industry Standards, where certain categories of service that provide access to specific types of material are also defined (e.g., 'high impact DIS' with the sole or predominant purpose of enabling access to 'high impact material') (see further discussion in chapter 8).

- **Accounting for emerging technological change**: One benefit of the co-regulatory industry codes scheme is that industry can leverage its forecasting experience and future product development and integration plans to identify emerging technological trends. This approach helps develop codes that are future-proof and responsive to those new technologies. However, in Phase 1, eSafety requested industry to resubmit the Search Engine services code because the earlier draft did not address risks associated with new technologies such as generative AI.

  eSafety encourages industry to account for new and emerging technologies relevant to the Phase 2 Codes and ensure that submitted codes deal with any associated issues.

- **Expression of matters in Notices to produce Codes**: The Phase 1 Notices contained matters for industry to address, which were distinct from the desired outcomes of Phase 1 as outlined in the September 2021 position paper. This separation led to measures being directed at outcomes rather than the specific matters in the Notices.

  In Phase 2, eSafety has attempted to define the matters as achievable outcomes that all industry sections can address. We discuss this further in chapter 8.

- **Industry steering/working group governance and processes**: From our preliminary engagements with industry, eSafety understands that Notice Recipients have considered how to streamline and improve the code drafting process for Phase 2. eSafety appreciates these continued efforts.

Phase 1 Codes and Standards

Consequently, six industry codes (collectively referred to as the **Phase 1 Codes**) addressing class 1A and class 1B material have now been registered. These six Phase 1 Codes (registered as Schedules to the Consolidated Head Terms) are now in effect. The **Phase 1 Industry Standards** (being the **Phase 1 RES Standard** and the **Phase 1 DIS Standard**) will

commence in December 2024. The Codes and Standards are available at the eSafety Register of Industry Codes and Standards for Online Safety.

This position paper seeks to build on these Phase 1 Codes and Standards, considering the different scope of matters addressed in the two phases. While leveraging the experiences of Phase 1, it acknowledges the need for different approaches and considerations for Phase 2.

5. Phase 2 notices to representative industry bodies

Notice recipients

On 1 July 2024, eSafety issued eight Notices under section 141(1) of the OSA (**the section 141 Notices**) to five industry bodies and associations (**the Notice Recipients**) for the development of the Phase 2 Codes:

- Digital Industry Group Inc. (**DIGI**)
- Communications Alliance Ltd (**CA**)
- Interactive Games and Entertainment Association (**IGEA**)
- Australian Mobile Telecommunications Association (**AMTA**)
- Consumer Electronic Suppliers Association (**CESA**).

In Phase 1, eSafety issued notices to these Notice Recipients as well as The Software Alliance (**BSA**). However, eSafety understands that BSA has withdrawn from the code-drafting process for Phase 2. Despite BSA's withdrawal, its membership base remains obligated to comply with any Codes or Standards produced from the Phase 2 process.

Before issuing the section 141 Notices, eSafety corresponded with the Notice Recipients about their capacity to represent each section of the online industry for Phase 2. The Notice Recipients indicated they are capable of representing the relevant sections of the online industry for developing the Phase 2 Codes.

BSA was identified as being one of the representatives of industry for the Hosting Services code in the Phase 1 Codes Head Terms. eSafety has received indications that members of the remaining Notice Recipients are still representative of the major Hosting Services providers. eSafety also notes that CA, which co-drafted the Phase 1 Hosting Code, remains an engaged member of the Notice Recipients. Further, the Notice Recipients have indicated that BSA and its members will be invited to contribute to the development of the Phase 2 Hosting Service Code.

eSafety also considers it crucial that parties representing the online pornography industry are adequately involved in the Phase 2 Code development process, to ensure their views are represented and to encourage compliance from those industry participants. There are multiple ways to achieve this.

As discussed in the Phase 1 position paper, eSafety recognises there is a continuum of representation.[24] On one end, industry participants can be consulted during the Code development process. On the other, they can be members of an industry association that represents their interests.

Accordingly, even if representatives of the online pornography industry – or any other relevant industry participants or community stakeholders - are not members of an industry association, eSafety expects industry to engage with them appropriately during this development process. To that end the Notice Recipients have informed eSafety that they have already contacted representatives of the online pornography industry, inviting them to participate in the development process and provide feedback. They have committed to giving them opportunities to be involved during this process. This is discussed further in chapter 10.

For the purposes of section 135 of the OSA, eSafety considers the Notice Recipients as representing the eight sections of the online industry that will need to develop Phase 2 Codes dealing with the matters explained below.

The Commissioner will reassess the adequacy of industry representation and consultation at the time any industry code is submitted for registration, as detailed in chapter 10.

Matters to be addressed in Phase 2 Codes

The Notices eSafety has issued under section 141 of the OSA request the Notice Recipients develop a Phase 2 Code or codes that deal with matters in similar terms to the following:[25]

1. Protect and prevent children in Australia from accessing or being exposed[26] to class 1C and class 2 material.
2. Provide end-users in Australia with effective information, tools and options to limit access and exposure to class 1C and class 2 material.

eSafety expects the industry codes will contain minimum compliance measures relevant to each section of the online industry's capacity to deal with these matters.

---

[24] eSafety Commissioner, *Development of Industry Codes: Position Paper*, September 2021, p. 57.

[25] The wording of matter 2 is amended slightly for the purpose of this Paper. In the s 141 Notices, it refers to the specific section of industry specified in the relevant Notice.

[26] During consultations conducted in the production of the Age Verification Roadmap, eSafety received feedback that the word 'exposure' should be avoided where possible due to the potential to stigmatise online pornography as something inherently negative (see p. 77 of the Age Verification Roadmap). eSafety acknowledges this feedback but notes that the word 'exposure' is language already used in the *Online Safety Act 2021*, particularly in s 108(4)(a) in respect of Restricted Access Systems. To ensure continuity between the aims of the RAS and the Phase 2 Codes we have adopted the same language in the matters to be addressed in the Codes.

To that end, this position paper sets out measures each section of industry can adopt in order to play its part in achieving the specified matters in the section 141 Notices across the technology stack. These measures build on the Phase 1 Codes and the measures outlined in eSafety's *Roadmap for Age Verification* and *Background Report*. This position paper also explains eSafety's approach to different types of class 2 material, highlighting those we believe warrant stronger measures, as well as those that are highly context-dependent, where scalable measures may pose greater challenges.

Timing

Under the OSA, the statutory minimum time for returning draft industry codes is 120 days (about four months). Industry feedback to eSafety has been that a minimum of 12 months is necessary to adequately develop a draft of Phase 2 Codes. eSafety acknowledges this input and believes a timeframe exceeding greater than the four-month statutory minimum but less than 12 months suggested by industry would be most appropriate in the circumstances.

As set out in chapter 4, during Phase 1, industry submitted the initial formally requested set of draft Codes to eSafety about 7 months after the issuing of the first section 141 Notices.

Drawing on the lessons learned from Phase 1, as well as the urgent need to establish protections for class 2 material online (particularly for children), eSafety has proposed a timeframe of **approximately 6 months** for the submission of Phase 2 Codes. eSafety considers this timeframe is reasonable and achievable as supported by several factors:

- eSafety has engaged with both the Notice Recipients and individual online industry participants for more than six months prior to the publication of this position paper. Those discussions about the approach to Phase 2 Codes prior to issuing the section 141 Notices included consultation on proposed scope, matters, measures, and timeframes, to support timely and effective codes development.
- eSafety has endeavoured to fulfil the co-regulatory aims of the OSA by proactively incorporating preliminary feedback from industry participants on the issues addressed in this position paper, where possible. This includes proposing a graduated scope for dealing with class 2 materials. This approach prioritises the most harmful content that Phase 1 did not cover, while also considering practical strategies for dealing with context-dependent materials.
- As in Phase 1, if industry submits draft Codes that eSafety considers could benefit from adjustments to ensure adequate community safeguards, eSafety will consider inviting Notice Recipients to submit revised codes.

- eSafety believes that foundational issues, including drafting principles and the general structure of the codes, can be adapted from Phase 1. This means less time is needed to consider these issues and the focus can be on the substantive provisions of the Phase 2 Codes.

Accordingly, eSafety has requested that Notice Recipients submit draft Codes to eSafety for final consideration by **19 December 2024**. Pursuant to section 141(5) of the OSA, eSafety has also requested a first draft of the Codes to be provided to eSafety for review by **3 October 2024** as an indicative target for achieving progress in the development of the codes.

If industry requires more time and believes the section 141 Notices should be varied to allow this, eSafety will consider such requests as they arise. As occurred in Phase 1, eSafety is open to considering requests for extensions of time.

6. Regulatory context

Relevant Australian developments

The positions and suggested approaches in this paper are supported by eSafety's *Roadmap for Age Verification*; recent eSafety research on young people's attitudes towards online pornography and age assurance;[27] and our *Age Assurance Issues Paper*. Addressing harms, particularly to young people, requires broad participation and support. We acknowledge that this work and the development of the codes align with other work such as eSafety's educational resources for young people, parents and educators. Additionally, cross-governmental initiatives to foster respectful relationships, efforts under the *National Plan to End Violence against Women and Children 2022-2032*, and the Government's upcoming Age Assurance Trial – encompassing age verification and age estimation technologies – are all complementary to Phase 2. Recognising this, eSafety has engaged with other national regulators and government departments with interconnected regulatory responsibilities about this paper and its suggested approaches.

eSafety also acknowledges recent public debate about the appropriate age for Australian children to access social media and other online services, including potential rules at the State/Territory or Commonwealth level to enact this.[28] The Phase 2 Codes are aligned with the NCS, focusing on age restrictions for accessing particular types of online material restricted to adults, rather than regulating access to online services based on age. However, these concepts converge in cases where a service is accessible by children and cannot ensure a low risk of encountering material likely to be classified as R18+ or above. eSafety will monitor developments related to minimum age requirements for access to services and contribute insights based on our regulatory remit, experience, and research. This will also be examined further in the DITRDCA's Age Assurance Trial (discussed below).

Statutory review of the Online Safety Act

Section 239A of the OSA requires an independent review of the OSA to begin within three years of it coming into effect (**the Review**). The Australian Government has appointed Ms Delia Rickard PSM to conduct the Review. The terms of reference include evaluating the effectiveness of statutory schemes in the OSA and assessing whether regulatory

[27] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023.
[28] Government of South Australia, *Nation-leading move to protect our children from social media*, 12 May 2024; Prime Minister of Australia, *Press conference – Sydney*, 14 June 2024; Queensland Government, *Call to balance social media impact on young Queenslanders*, 21 May 2024.

arrangements should be amended for the Online Content Scheme, including the legislative framework governing industry codes and standards.

An Issues Paper, released in April 2024, seeks public input on a range of issues, including the OSA's role in restricting children's access to age-appropriate content through age assurance technologies and the Commissioner's authority to address access to violent pornography.

The final report will be provided to the Minister for Communications by 31 October 2024.

During preliminary discussions on Phase 2 Codes, some industry participants suggested delaying code development until the OSA review concludes. They argue that the effectiveness and regulatory arrangements of industry codes provisions will likely be considered in the review.

However, eSafety believes delaying the Phase 2 Codes would not be in the best interests of children or the broader Australian community. Legislative reforms stemming from the review could take a significant amount of time to materialise. Starting development of Phase 2 Codes now will provide, at the very least, essential interim protections relating to class 2 material. eSafety therefore views the prompt implementation of Phase 2 Codes as urgent and in the community's interest.

Review of the Privacy Act and Australian Children's Online Privacy Code

The Attorney-General's Department (AGD) began reviewing the Privacy Act in October 2020, following recommendations from the Australian Competition and Consumer Commission's 2019 Digital Platforms Inquiry. The review led to 116 proposed reforms aimed at enhancing privacy protections for Australians and aligning Australia's privacy laws with global standards and enhancing protections for Australians' privacy. On 28 September 2023, the Australian Government released its response to the Privacy Act Review Report.

Of note, the Government agreed in principle to the following recommendations:

- **Broadened definition of personal information:** Personal information should include technical and inferred information (such as IP addresses and device identifiers) if it can identify an individual (proposal 4.1). The definition of 'collection' should also be amended to cover information obtained from any source and by any means, including inferred or generated information (proposal 4.3)
- **Mandatory privacy impact assessments:** Non-government entities should be required to conduct Privacy Impact Assessments for activities with high privacy

risks (proposal 13.1). Enhanced risk assessments should be considered for biometric information (proposal 13.2)

The Government also agreed or agreed in principle to several protections specific to children:

- **Best interests of the child:** Entities must consider the best interests of the child when determining if collection, use, or disclosure of personal information is 'fair and reasonable in the circumstances', balanced against other factors, including the interests of other children (proposal 16.4).
- **Clear and understandable notices:** Entities must provide privacy notices and policies that are clear and understandable, especially when specifically addressing children. (proposal 16.3).
- **Prohibition on targeting children**: Targeting[29] children should be prohibited except for targeting that is in the best interests of the child (proposal 20.6). This includes where platforms may target content to prevent children from seeing age-sensitive advertisements.
- **Valid consent:** The Privacy Act should codify that valid consent must be given with capacity (proposal 16.2). Exceptions should be made where parental or guardian involvement could harm the child.
- **Ban on direct marketing to children:** Entities should be prohibited from direct marketing to a child (proposal 20.5) unless the personal information was collected directly from the child and the direct marketing is in the child's best interests.
- **Prohibition on trading children's personal information:** Entities should be prohibited from trading the personal information of children (proposal 20.7).
- **Children's Online Privacy Code:** The Government also agreed that a **Children's Online Privacy Code** should be developed that applies to online services that 'are likely to be accessed by children' (proposal 16.5).

The Children's Online Privacy Code is expected to clarify the principles-based requirements of the Privacy Act in more prescriptive terms and provide guidance on upholding the best interests of the child in the design of online services. This will include how to meet requirements regarding targeting, direct marketing and trading data.

[29] The proposed definition of 'targeting' in the Review Report is to 'capture the collection, use or disclosure of information which relates to an individual including personal information, deidentified information, and unidentified information (internet history/tracking etc.) for tailoring services, content, information, advertisements or offers provided to or withheld from an individual (either on their own, or as a member of some group or class)' (proposal 20.1).

The Code developer will determine the specific requirements of the Code. However, the Code may cover:[30]

- assessing a child's capacity and establishing their age
- limiting certain collections, uses and disclosures of children's personal information
- default privacy settings
- enabling children to exercise privacy rights
- balancing parental controls with a child's right to autonomy and privacy.

To meet these requirements, the Government expects the Code may also address whether entities need to take reasonable steps to establish an individual's age with a level of certainty that is appropriate to the risks, such as by implementing age assurance. Enhanced legislated privacy protections proposed in the review will also impact the implementation and use of age assurance by organisations and how that information may be used.

The Government aims for the Children's Online Privacy Code to align with the UK *Age Appropriate Design Code*, balancing child safety with preserving online privacy for children. The interaction of the UK *Age Appropriate Design Code* with age assurance measures is discussed in further detail in chapter 8 below.

The Government response noted that the developer of the Children's Online Privacy Code should consult broadly, including with children, parents, child development experts, child welfare advocates, industry and eSafety. eSafety continues to engage with the Office of the Australian Information Commissioner (**OAIC**) on regulatory developments, both bilaterally and through the Australian Digital Platform Regulators Forum (**DP-REG**).

Considering the potential intersections between the Phase 2 Codes and a Children's Online Privacy Code in Australia, eSafety finds it useful to examine how relevant privacy and safety regulators have worked together in the UK. The UK Age Appropriate Design Code is administered by the Information Commissioner's Office (**ICO**), which works closely with Ofcom, the online safety regulator. Ofcom and ICO have a memorandum of understanding (**MoU**) on online safety and data protection, enabling information sharing and collaboration to maximise coherence.[31]

Under the Digital Regulation Co-operation Forum, ICO and Ofcom have jointly commissioned research on families' attitudes to age assurance in the UK and ways to

---

[30] Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Australian Government response to the Roadmap for Age Verification*, 30 August 2023.
[31] Information Commissioner's Office, *Memorandum of Understanding between the Information Commissioner and Ofcom clause 11*, Ofcom website, 2019.

measure the accuracy of different age assurance solutions.[32] Since the release of the ICO's Opinion on Age Assurance, the ICO and Ofcom have continued to engage to ensure regulatory alignment between the UK Age Appropriate Design Code and the UK *Online Safety Act 2023*. In May 2024, the ICO and Ofcom published a joint statement on collaboration in regulating online services, building on their 2022 statement on online safety and data protection.[33] Age assurance is listed in this statement as an important collaboration theme for the two agencies, enabling new ways of collaboration and information sharing when engaging with companies of mutual interest.[34]

Aligned approaches to shared issues such as age assurance can support regulatory coherence and reduce the regulatory burden for industry while ensuring measures support, rather than undermine, the balanced objectives of data protection and online safety.

eSafety views the working relationship between ICO and Ofcom as a leading example of privacy and online safety regulators working together to protect children online. eSafety will work closely with the OAIC through DP-REG to ensure industry codes enhance safety and preserve privacy. eSafety will also engage in other coordinated efforts, such as through the Global Online Safety Regulators Network, to further these aims.

National Classification Scheme reforms

A two-staged reform process to modernise the National Classification Scheme is currently underway. Legislation to implement parts of the stage 1 reforms commenced on 14 March 2024. Notably, this legislation includes the addition of simulated gambling as R18+ content in the *Computer Games Guidelines,* which is relevant for the Phase 2 Codes.

Public consultation on stage 2 reforms concluded in May 2024. Relevant aspects of this consultation include:

- establishing an independent Classification Advisory Panel or similar body to advise Commonwealth, State and Territory governments on possible updates to classification criteria
- establishing a single national regulator responsible for classification at the Commonwealth level

[32] Information Commissioner's Office, *Age Assurance research*, ICO website.
[33] Ofcom and the Information Commissioner's Office, *Online safety and data protection: A joint statement by Ofcom and the Information Commissioner's Office*, 25 November 2022, Information Commissioner's Office, *A Joint Statement by Ofcom and the Information Commissioners Office on collaboration on the regulation of online services*, 1 May 2024.
[34] Information Commissioner's Office, *A Joint Statement by Ofcom and the Information Commissioners Office on collaboration on the regulation of online services*, 1 May 2024.

- reviewing regulatory and governance arrangements for classification.[35]

At this stage, none of the proposed reforms seek to modify any relevant definitions relating to class 2 material within the Classification Guidelines.

As in Phase 1, the scope of content for the Phase 2 Codes is based on existing definitions of class 2 material as drawn from the Classification Guidelines. However, class 2 material relates to a broader range of content than that covered in Phase 1. Specifically, class 2 material may include high-impact depictions of 'themes,' which can be highly context-dependent.

With this in mind, eSafety has proposed industry adopt a graduated approach to measures across the spectrum of class 2 material. This approach also aims to align with proposed measures required in international jurisdictions that deal with the same or similar material, as much as possible.

Age Verification Roadmap

In March 2023, eSafety published its Age Verification Roadmap, examining the feasibility and implementation of a mandatory age verification mechanism in Australia. This roadmap provided a holistic overview of approaches to address the risks and harms associated with children's access to online pornography. In August 2023, eSafety supplemented this with its Age Verification Roadmap Background Report. To inform the Roadmap and Background Report, eSafety:[36]

- conducted a call for evidence and held extensive multi-sector consultations
- undertook both desktop and primary research, including a survey and focus groups with participants aged 16-18, supported by further discussions with the eSafety Youth Council
- commissioned an independent assessment of age assurance and online safety technologies
- consulted with relevant agencies and departments across government.

Importantly, throughout the development of the Age Verification Roadmap, eSafety took a human rights-based approach, considering the rights, best interests and evolving capacities of children, as well as the rights of parents, carers, and other adults, including sex workers and performers and producers of online pornography.[37] Consequently, eSafety endorsed a proportionate and balanced approach to restricting children's access to online pornography.

[35] Department of Infrastructure, Transport, Regional Development, Communications and the Arts, *Modernising Australia's National Classification Scheme - Stage 2 Reforms*.
[36] eSafety Commissioner, *Roadmap for Age Verification*, March 2023, p. 5.
[37] eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023, p. 43.

This approach aligns with the United Nations Committee on the Rights of the Child, supporting the child's best interests while also respecting the rights of adults to consume and produce pornography in a safe and lawful manner.[38]

In the Age Verification Roadmap, eSafety recommended that the Australian Government develop, implement, and evaluate a cross-government trial of age assurance technologies in Australia before mandating their use. eSafety also recommended other complementary measures that could inform and be adopted in the Phase 2 Codes either before or during this trial.

Age Assurance Trial

In May 2024, the Federal Government announced an Age Assurance Trial to test the efficacy of age verification and age estimation technologies in protecting children from encountering online pornography and other high-impact online content.[39] The Trial will also consider consultation and research on age limits for access to social media.[40]

The DITRDCA will lead the Age Assurance Trial, supported by a cross-government working group. The trial will run in parallel with the Phase 2 Codes development over the remainder of 2024. These processes are expected to be mutually reinforcing, with the overall goal of improving the online safety experiences of Australians.

The Age Assurance Trial will not specifically inform the creation of the Phase 2 Codes. Instead, eSafety will be a member of the cross-government working group and will consider how outputs from the trial can inform and support our work during the development of the Phase 2 Codes. For example, while the Phase 2 Codes may outline *what* industry is expected to do (e.g., take reasonable and appropriate steps to confirm users' age) and *when* or *where* they should do it, the outcomes of the Age Assurance Trial may support industry with the *how* (e.g., by informing what reasonable and appropriate steps for compliance may be best within the Australian context).

The research and evidence base about methods of age assurance and age verification technologies is consistently evolving. The eSafety *Age Assurance Issues Paper* considers recent developments in this field, building on the work of the *Age Verification Roadmap*. However, the task of determining which specific methods of age assurance or verification should be adopted in Australia will be conducted through the DITRDCA's Age Assurance Trial. Therefore, no position on specific methods is provided in this paper.

---

[38] eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023, p. 36.
[39] Attorney-General's Portfolio, *Tackling Online Harms*, 1 May 2024; Minister for Communication, *Boosting connectivity and safety for Australians*, 14 May 2024.
[40] Australian Parliament, Environment and Communications Legislation Committee, *Estimates*, 30 May 2024, Department of Infrastructure, Transport, Regional Development, Communications and the Arts.

Instead, eSafety proposes that industry should establish compliance measures focused on establishing expectations around age assurance in the Phase 2 Codes. These measures should set out the base expectations for what any adopted age assurance measures should achieve. Ideally, these should align with the existing expectations and obligations under complementary schemes, as set out below.

Australia's Digital ID System

The Australian Government has passed legislation to enhance and expand the existing Australian Digital ID system. Digital IDs are already offered to Australians through myGovID.

The *Digital ID Act 2024* (Cth) will commence on 1 December 2024. It will allow for the expansion of the current Australian Government Digital ID System for use by the Commonwealth, state and territory governments and eventually private sector organisations.[41]

The Digital ID System is being expanded to address the increasing shift to digital transactions and online engagement.[42]

The Australian Competition and Consumer Commission (**ACCC**) will be the Digital ID Regulator, while the Office of the Australian Information Commissioner (**OAIC**), will regulate the privacy aspects of Australia's Digital ID System.

The adoption of the Digital ID System in Australia and similar systems across the world[43] demonstrates how a broader shift towards online identity verification is occurring and becoming normalised, applying to user interactions which go beyond the age assurance or verification context.

Complementary eSafety Schemes

Restricted Access Scheme

A Restricted Access System is an access-control system that meets the requirements under the *Online Safety (Restricted Access Systems) Declaration 2022* (Cth) (**RAS Declaration**). This sets out the minimum requirements for access-control systems used by social media services, relevant electronic services and designated internet services provided

[41] Australian Government, *Digital ID Legislation*, Australia's Digital ID System.
[42] Australian Government, *About Digital ID*, Australia's Digital ID System.
[43] For example, Singapore's Singpass Digital ID system (see singpass.gov.sg/main), Sweden's BankID system (see norden.org/en/info-norden/electronic-identification-sweden) and India's Aadhaar system (uidai.gov.in/en).

from Australia. Its primary aim is to restrict children's access to R18+ content (i.e. a subset of class 2 material) online, upon receiving a notice from eSafety.

Rather than mandating specific technologies or processes, the RAS Declaration states that an access-control system must:

- require an application be made by a person to access the relevant material, declaring they are at least 18
- incorporate reasonable steps to confirm an applicant is at least 18
- give warnings about the nature of the material and safety information about how a parent or guardian may control access to the material
- limit access to the material unless certain steps are followed.

Accordingly, the RAS Declaration takes a non-prescriptive approach. As explained in the Explanatory Statement to the RAS Declaration, this was done to:[44]

- protect children from exposure to unsuitable content without imposing unreasonable financial or administrative burdens on service providers
- give service providers flexibility to implement restriction measures that reflect the nature of their services, their business models, and their size, capacity, capability and maturity
- allow for advances in relevant technologies.

Several organisations have adjusted their practices in response to the RAS Declaration. For example, Google has introduced age assurance steps on YouTube and Google Play, which require some Australian users to provide additional proof of age when attempting to watch mature content on YouTube or downloading content on Google Play.[45]

Given that many social media services, relevant electronic services, and designated internet services subject to Phase 2 Codes must already comply with the requirements of the RAS Declaration, eSafety believes these requirements could serve as foundational elements on which Phase 2 Codes measures could be built. This is explored in more detail in chapter 8.

Basic Online Safety Expectations

The Basic Online Safety Expectations apply to social media services, relevant electronic services, and designated internet services.

---

[44] *Explanatory Statement for the Online Safety (Restricted Access Systems) Declaration 2022* (Cth).
[45] Google Australia, *Ensuring Age Appropriate Experiences*, Google website, 2022.

The *Online Safety (Basic Online Safety Expectations) Determination 2022* (**BOSE Determination**) effective from 23 January 2022 and amended by the *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* on 30 May 2024, sets out clear expectations that these services take reasonable steps to protect Australians from unlawful and harmful material and activity that falls within the remit of the OSA or impedes the online safety of Australians. Under the OSA, eSafety has the authority to request information from online service providers about their compliance with the expectations. This is intended to promote transparency and accountability from industry and incentivise improvements in safety standards.

The BOSE Determination sets out non-enforceable expectations for social media services, relevant electronic services, and designated internet services. These expectations include:

- **Child protection measures:** Take reasonable steps to ensure technological or other measures are in place to prevent children from accessing class 2 material. Examples include age assurance mechanisms, child safety risk assessments, and implementing improved technologies and processes for preventing access by children to class 2 material[46]
- **Reporting mechanisms:** Provide clear and readily identifiable reporting mechanisms that allow end-users to report, and make complaints about, class 2 material and to review and respond to those reports within a reasonable period[47]
- **Enforcement:** Take reasonable steps (including proactive steps) to detect breaches of terms of use, policies and procedures and standards of conduct, and ensure that any penalties specified for breaches of its terms of use, policies and procedures in relation to the safety of end-users, and standards of conduct for end-users, are enforced against all accounts held or created by the end-user who breached the terms of use and, where applicable, breached the policies and procedures, and standards of conduct, of the service[48]
- **Terms of use:** Ensure that the service has clear and readily identifiable mechanisms that enable end-users to report, and make complaints about, breaches of the service's terms of use[49]
- **Child-centric design:** Ensure that the best interests of the child guides the design and operation of any service that is likely to be accessed by children[50]

---

[46] *Online Safety (Basic Online Safety Expectations) Determination 2022* (Cth), s 12.
[47] *Online Safety (Basic Online Safety Expectations) Determination 2022* (Cth), s 13.
[48] *Online Safety (Basic Online Safety Expectations) Determination 2022* (Cth), s 14(2); *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* (Cth), schedule 1.
[49] *Online Safety (Basic Online Safety Expectations) Determination 2022* (Cth), s 15.
[50] *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* (Cth), schedule 1.

- **User autonomy:** Take reasonable steps to provide controls that give end-users the choice and autonomy to support safe online interactions[51]
- **Safe recommender systems:** If a service uses recommender systems, take reasonable steps to consider end-user safety and incorporate safety measures in the design, implementation and maintenance of recommender systems on the service; and proactively minimise the extent to which recommender systems amplify material or activity on the service that is unlawful or harmful[52]

In general, Industry Codes and Standards impose narrower obligations for RES, DIS and SMS compared to the BOSE Determination because they are mandatory and enforceable, and focus solely on class 1 and 2 materials. In contrast, the Determination's expectations are broader, encompassing unlawful and harmful material and activity, though they are not enforceable.

However, there are instances where specific mandatory measures in Codes and Standards for addressing class 1 or class 2 material may directly correspond to expectations outlined in the BOSE Determination.

Accordingly, eSafety suggests in chapter 8 that industry consider aligning measures in the Phase 2 Codes with expectations set out in the BOSE Determination where it makes sense, particularly in respect of applying appropriate age assurance to limit access to class 2 material.

International regulatory landscape

Countries around the world are introducing legislation and regulation to combat online harms. While each nation operates within different political and cultural contexts and norms, online safety frameworks and regulatory mandates share several key elements. Beyond state-driven efforts, the online industry has also taken voluntary steps to improve practices.

Preventing children's access and exposure to harmful content is a primary goal of online safety regulations in many countries, including Australia, with a focus on the role of age assurance. There is also a growing focus on regulations that promote user empowerment – one of the principles of Safety by Design.

eSafety has aligned the proposed measures for Phase 2 Codes with international standards to address similar online content and its associated harms. eSafety's commitment to

[51] *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* (Cth), schedule 1.
[52] *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* (Cth), schedule 1.

regulatory coherence is also set out in the Global Online Safety Regulators Network's position statement, published in May 2024.

International developments in age assurance are also explored in eSafety's Age Verification Roadmap and Background Report, and more recently, in our Age Assurance Issues Paper.[53]

Industry associations developing codes under the OSA should remain aware of international approaches and aim to achieve a consistent global approach, while still respecting regional differences.

[53] eSafety Commissioner, *Roadmap for Age Verification*, March 2023; eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023; the eSafety *Age Assurance Issues Paper*.

Table 1 – International regulatory approaches

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
Ireland – *Online Safety and Media Regulation Act 2022*	*Online Safety Code for VSPSs* Includes protecting children from harmful content, such as: • Pornography • Promotion of eating disorders Teaching/promotion of self-harm or suicide (or dangerous behaviours prejudicial to safety) [54]	• Terms and conditions (T&Cs) that prohibit the uploading and sharing of harmful content, to rate adult content as not suitable for children, and to preclude children from services that provide adult-only content (e.g., pornography and gratuitous violence). • T&Cs enforced through content and account detection, and account removal. • Using effective age assurance measures to ensure that 'adult-only' content cannot be seen by children.[55] • Parental control measures when platforms permit users under 16 years. • Accessible and user-friendly reporting, and content rating systems. • Effective procedures for handling user complaints. • Measures to promote media literacy and provide online safety info to users.	• Designated video-sharing Platform Services (VSPS) under the jurisdiction of Ireland.[56]	Expected to come into force in Q4 2024.[57]

[54] The draft code also includes general measures for all users in relation to content that would be considered Class 1 content in Australia. The code also includes protecting children from bullying and humiliation, and requirements around commercial communications, which are outside the scope of Class 2 material in Australia. See Draft Online Safety Code, Part A, 10.1 a-c. cnam.ie/coimisiun-na-mean-to-notify-online-safety-code-to-european-commission.

[55] An age assurance measure based solely on self-declaration of age by users of the service is not considered an effective measure, See Draft Online Safety Code, Part B, 12.11. cnam.ie/coimisiun-na-mean-to-notify-online-safety-code-to-european-commission.

[56] As of January 2024, this includes Facebook, Instagram, YouTube, TikTok, LinkedIn, X, Pinterest, Tumblr, Reddit and Udemy. cnam.ie/designation-notices/

[57] On 27 May 2024, the draft Code was submitted to the European Commission under the TRIS Directive process for a period of consultation prior to assent. If there are no objections, the Code can be adopted into law after the consultation expires (estimated at 3-4 months).

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
United Kingdom *Online Safety Act 2023*	*Protection of Children Codes of Practice*[58] Requirements for Primary Priority Content, which includes: • Pornography Content that encourages, promotes, or instructs suicide, self-harm, and eating disorders.[59]	• Governance requirements including annual risk assessments and mitigations, training requirements for Trust and Safety (T&S) workforce and content moderators and tracking for new and emergent harms to children. • Content moderation including removal and demotion, and AI tools when services have sufficient resources and expertise. • Responding to user reports and having easy to use complaints systems. • User controls and empowerment measures including ability to block accounts and demote and restrict types of content and for search services, safe search features, and removal of predictive search suggestions, and provision of crisis information (relating to self-harm, eating disorder, and suicide). • Terms of service must deal with content harmful to children and must be accessible, easy to locate and consistently enforced. • Implement 'highly effective age assurance'[60] with extra technical requirements for services whose primary purpose is to provide pornographic content. • In-service support and information for children, and signposting/referral for users who demonstrate risk signals. • Additional obligations for recommender systems including demotion of potentially harmful content.	• User-to-user services, which includes social media video-sharing, private messaging, dating services, gaming services, file- and audio-sharing services.[61] • Search engine services.	Draft code currently under public consultation. Final code expected to come into force in Q1 2025 pending parliamentary approval.

[58] The Children's Safety Code relates to Matter 1 in Australia (children's access and exposure). Requirements related to Matter 2 in Australia (tools and information for all end-users), will be considered in a further code that will cover 'platforms transparency and user empowerment' for larger platforms and services. Ofcom has also drafted an Illegal Content Code of Practice.

[59] The code also includes content which is outside the scope of class 2 material in Australia, such as abuse, content that incites hatred or bullying, violence, dangerous stunts and challenges, and harmful substances.

[60] 'Highly effective age assurance' – must meet the required thresholds for: technical accuracy (success rate), robustness (success rate and accuracy in real world context), reliability (consistency of outputs), and fairness (extent to which biases and discrimination is minimised). See Ofcom, *Protecting children from harms online*, Annexes 10-15, A10.

[61] See tabs on difference types over services, Ofcom, *Online safety rules: what you need to know*, 26 October 2023.

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
Singapore *Online Safety (Miscellaneous Amendments) Act, updating the Broadcasting Act 1994*	*Code of Practice for Online Safety* Services are required to enhance online user safety, particularly for children, and curb the spread of harmful content on their service, including: • Sexual content • Violent content • Suicide and self-harm content Content facilitating vice and organised crime.[62]	• Publish community guidelines and standards that address harmful content. • Provide tools that restrict the visibility of harmful content. • Prohibit the targeting of harmful content to children including advertisements, promotions and recommendations. • Provide children and their parents/guardians tools to manage their safety and minimise exposure to harmful content. • Provide access to support services for end-users who use high-risk search terms (e.g., terms relating to self-harm and suicide). • Provide education and awareness raising programs and initiatives to protect children from harmful and/or inappropriate content. • Require differentiated accounts with default settings for children, with clear warnings of implications if the user opts out. • User reporting and removal or restriction of harmful content. • Transparency reporting.	• Designated social media services.[63] • Note: Singapore is currently consulting on a Code of Practice for App Distribution Services, to be published late-2024 or early-2025.	In effect from 18 July 2023.

[62] The Code of Practice includes general measures for all users, as well as specific measures for children, and covers protecting all end-users from harm in relation to content that would be considered Class 1 content in Australia. The code also includes Cyberbullying content and Content endangering public health as categories of harmful content, which is outside the scope of Class 2 material in Australia. See: Guidelines on categories of harmful content on imda.gov.sg/regulations-and-licensing-listing/content-standards-and-classification/standards-and-classification/internet/online-safety.

[63] As of 18 July 2023, this includes Facebook, HardwareZone, Instagram, TikTok, X and YouTube. See: imda.gov.sg/regulations-and-licensing-listing/content-standards-and-classification/standards-and-classification/internet/online-safety.

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
France *Loi Visant à Sécuriser et à Réguler l'Espace Numérique*	Reference framework to prevent children's access to pornography	• Age-verification.[64] • Requiring measures to prevent exposure, for example blurring webpages.	• Websites. • Intermediary services.	Expected to come into force in Q3 2024.
France[#] *Loi Visant à renforcer le contrôle parental sur les moyens d'accès à internet*	Preventing access to content likely to harm the physical, mental, or moral development of children[65]	• Provide and activate parental controls that prohibit children from downloading or accessing prohibited content, e.g., pornography, violence, discrimination/hate. • Implementation of local, on device parental controls that do not otherwise process personal data of the child. • Provision of information in relation to parental controls and exposure to inappropriate content.	• Equipment providers and manufacturers.	Expected to come into force from 13 July 2024.

[#] This table was updated on 6 August 2024 due to a data error on the France and European Union content.

[64] Under the reference framework, French regulator Arcom will set out technical requirements relating to the reliability of age checks on users and respect for their privacy, and specify the procedures for carrying out and publicising audits. A draft framework on technical requirements for age-verification under the SREN was published in April 2024, see arcom.fr/vos-services-par-media/consultations-publiques/consultation-publique-sur-le-projet-de-referentiel-determinant-les-exigences-techniques-minimales-applicables-aux-systemes-de-verification-de-lage-mis-en-place-pour-acces-contenus-pornographiques-en-ligne; LOI n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique (1) - legifrance.gouv.fr/jorf/id/JORFTEXT000049563368 Note: these websites are published in French – any quotations in English in this position have been produced using a translation tool.

[65] LOI n° 2022-300 du 2 mars 2022 visant à renforcer le contrôle parental sur les moyens d'accès à internet (1) – See legifrance.gouv.fr/jorf/id/JORFTEXT000045287677/, article 1; National Frequency Agency (ANFR), *Publication of the inventory of parental control devices with regard to the regulatory provisions applicable in July 2024*, 21 March 2024. Note: these websites are published in French – any quotations in English in this position have been produced using a translation tool.

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
European Union[##] *Digital Services Act*	• Content that may impair minors' health, physical, mental, and moral development[66] • Requirements for platforms that are primarily used for the dissemination to the public of pornographic content[67]	• Risk management framework, assessments, and mitigations, include taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate;[68] • Governance requirements such as compliance officers.[69] • Age assurance to prevent children from accessing age-inappropriate content.[70] • Measures to alter results returned by recommender systems, such as demotion.[71] • User empowerment tools for young people.[72] • Accessible, 'easily understandable' and enforced T&Cs.[73] • Transparency reporting.[74] • Default design to the highest-level privacy, safety, and security for minors.[75]	• VLOPs and VLOSEs[76]. • Online platforms (including gaming with user-to-user communications). • Hosting services. • Intermediary Services (including app stores and marketplaces).	DSA came into full force on 17 February 2024. The EU's voluntary Code of Conduct for age-appropriate design is in development.[77]

[##] This table was updated on 6 August 2024 due to a data error on the France and European Union content.
[66] Recital 81 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services; While the DSA does not explicitly name harmful content types such as content that depicts or promotes suicide, self-harm, or eating disorders, recent enforcement and information gathering actions taken by the European Commission highlights the intention for this category of risk to include such harms.
[67] Recital 87, ibid
[68] Articles 34 & 35 ibid
[69] Article 41, ibid
[70] Recital 71, ibid
[71] Recital 88 ibid
[72] Recital 40, ibid
[73] Recital 46, ibid
[74] Recital 65, ibid
[75] Recital 71, ibid
[76] Very large online platforms (VLOPs) and very large online search engines (VLOSEs) are regulated by the European Commission. In December 2023, the European Commission designated major pornography sites Pornhub, Stripchat and Xvideos as VLOPs. Other services are regulated by the Digital Services Coordinated in the country where the platform is domiciled.
[77] See Article 45, Digital Services Act. Note: Codes are non-binding under the DSA.

Country / Regulation	Matters /content covered	Measures/requirements of industry	Industry sections	Timing
Adult Content Standards in Combating Online Child Sexual Abuse Imagery[78]	Directed at guiding the online adult industry to address CSAM	• Subject anyone visiting, publishing, or appearing in material on an adult services platform to age verification measures to confirm they are over 18 years old at the time of production before content is permitted to be published. • When required by law, adult services must display an age disclaimer asking the visitor to confirm that they are 18 years or older, on the landing page or any direct URL link from a search engine before revealing any adult material. • Adult services should include information on how parents can enable controls in popular browsers and operating systems to prevent children from accessing the platform.	• Voluntary code, developed between Aylo (parent company of Pornhub) and the Internet Watch Foundation (IWF); only online adult content providers seeking full IWF membership must comply with code.	In force from May 2024.

[78] Internet Watch Foundation, *Effectively Tackling Child Sexual Abuse Online: A Standard of Good Practice for Adult Services*, May 2024; Internet Watch Foundation, *Effectively Tackling Child Sexual Abuse Online: A Standard of Good Practice for Adult Services – Supporting Document*, May 2024.

7. Scope of the Phase 2 Codes

The September 2021 position paper set out the types of materials to be captured in the Phase 2 Codes as follows:

Phase	Class Subcategory	Material	National Classification Scheme
Phase 2	Class 1C	• Online pornography – material that describes or depicts specific fetish practices or fantasies	• Class 1 • Refused Classification
Phase 2	Class 2A	• Online pornography – other sexually explicit material that depicts actual (not simulated) sex between consenting adults	• Class 2 • X18+ • Category 2 restricted
Phase 2	Class 2B	• Online pornography – material which includes realistically simulated sexual activity between adults. Material which includes high-impact[79] nudity • Other high-impact material which includes high-impact sex, nudity, violence, drug use, language and themes. 'Themes' includes social Issues such as crime, suicide, drug and alcohol dependency, death, serious illness, family breakdown and racism • Simulated gambling in computer games (from September 2024)[80]	• Class 2 • R18+ • Category 1 restricted

Subsequently, and relevantly for the Phase 2 Codes, industry adopted slightly altered definitions of 'particular online pornography' and 'online pornography' in the Head Terms to the Phase 1 Codes:

> **Particular online pornography**
>
> *Class 1 material that without justification:*
>
> - *depicts, expresses or otherwise deals with matters of sex in such a way that the material offends against the standards of morality, decency and propriety generally accepted by reasonable adults to the extent that such material should not be classified; or*

[79] Impact may be higher where content is detailed, accentuated, or uses special effects, prolonged, repeated frequently, realistic or encourages interactivity.

[80] Actual gambling is addressed and regulated by the ACMA under the *Interactive Gambling Act 2001* (Cth), and by state and Commonwealth governments and agencies under the National Consumer Protection Framework for Online Wagering.

- *includes or contains gratuitous, exploitative or offensive depictions of*
  - *(i) sexual activity accompanied by fetishes or practices which are offensive or abhorrent; or*
  - *(ii) incest fantasies or other fantasies which are offensive or abhorrent.*

Note: Fetish means an object, an action or a non-sexual part of the body which gives sexual gratification

Online pornography

Material depicting sexual activity between or in the presence of only consenting adults, or material depicting the nudity of only consenting adults.

eSafety considers both definitions can be adopted, or adapted with slight amendments, for Phase 2. We set out further details on suggested adaptions at **Table 2.**

Suggested scope of class 2 material to be dealt with by Phase 2 Codes

To determine what materials are class 2, the National Classification Board provides guidance based on:

- sections 9A and 11 of the Classification Act
- the National Classification Code
- the Classification Guidelines.

The National Classification Code sets out that any classification decision should adhere to the following principles:

- Adults should be able to read, hear and see what they want.
- Minors should be protected from material likely to harm or disturb them.
- Everyone should be protected from exposure to unsolicited material that they find offensive.
- The need to take account of community concerns about:
  - depictions that condone or incite violence, particularly sexual violence
  - the portrayal of persons in a demeaning manner.

Each set of Classification Guidelines sets out several classifiable elements and explains how their impact can determine the classification of material. Broadly, these elements are:

- Violence
- Sex
- Language
- Drug use
- Nudity
- Themes.

In eSafety's consultation with industry prior to the release of this position paper, industry expressed concern that protective measures might not scale effectively for all class 2 material. This is because the impact of material – and therefore its classification – often depends on its context. For example, violent imagery in a news article about a current global event might be classified differently due to its context, reducing its impact.

eSafety will, when approaching different types of class 2 material:

- focus on materials with a well-established evidence base of associated harms
- consider the tools available to industry for applying measures at scale and where industry should commit to developing tools
- consider international developments that seek to address similar materials and harms, aiming for a consistent global approach while still respecting regional differences.

Industry has taken a practical approach to dealing with the different sub-categories of harmful materials in Phase 1. For example, in the Phase 1 Code for social media services, industry developed distinct minimum compliance measures to deal with the spread of child sexual abuse material (**CSAM**) and pro-terror material on their platforms. Specific technologies known to help identify CSAM, such as Photo DNA and CSAI Match, were identified in the minimum compliance measures addressing CSAM. Meanwhile, key word searches were more relevant for identifying pro-terror material.

eSafety suggests industry take a similar approach for Phase 2 and develop different measures for different types of material under class 2. eSafety suggests the codes should focus on addressing harms in the following key areas with reference to suggested considerations (see below at Table 2). To summarise, eSafety proposes that industry should:

Prevent children's access to online pornography and high-impact nudity:

- Implement measures to block children's access to particular online pornography, online pornography and high-impact nudity.

- Build on the Phase 1 Codes head terms definition of 'online pornography' with minor adjustments to include realistically simulated, generated and animated sexual content; high-impact text-based sexual content, including interactive services such as chatbots and AI models providing pornographic content; and high-impact nudity, ensuring protections are clear.

Address high-impact violence, drug use and themes

- Build on measures from Phase 1 Codes and Standards which already cover the most harmful crime, violence and drug-related material (class 1B).
- Implement additional Phase 2 measures, such as effective policies on services likely accessed by young children, to prevent exposure to age-inappropriate content involving crime, violence or drugs which would not be picked up by the Phase 1 measures.

Apply graduated risk-based protections for remaining classifiable elements and themes in class 2

- Align protections for themes of suicide and serious illness with international requirements where possible.
- Manage risks of children's exposure to simulated gambling in video games by enforcing access policies where there is a high risk.
- Maintain and enforce policies about other themes and classifiable elements on services likely to accessed by young children.
- Provide all users, including children, young people, parents/carers, and adults, with effective tools, information, and options to minimise exposure to class 2 material.

Table 2 – Suggested scope for Phase 2 Codes

Type of material	Classifications	Suggested Considerations for proposed scope
Online pornography and high impact nudity	Class 1C – material that describes or depicts specific fetish practices or fantasies (RC). Class 2A – Other sexually explicit material that depicts actual (not simulated) sex between consenting adults (X18+ / Cat 2). Class 2B – Material which includes realistically simulated* sexual activity between adults and material which includes high impact nudity (R18+ / Cat 1).	**The priority focus of the Phase 2 Codes should be on preventing children's access to pornography.** *Note: Sexual violence material falls within class 1A material, and should be dealt with by industry according to the relevant provisions of the Phase 1 Codes and Standards.* We suggest **all services need to contribute proactively** to prevent and address harms related to this material, through a range of measures which complement and work alongside age-assurance systems which we suggest could be applied at the account level to flow across a given product and service ecosystem. In the Phase 1 Codes head terms, industry defined 'particular online pornography' which pertains to class 1C material. This definition can continue to be used in Phase 2. In the head terms, industry also defined 'online pornography' as 'material depicting sexual activity between or in the presence of only consenting adults, or material depicting the nudity of only consenting adults'. We consider this was a workable definition for Phase 1, and should be built on to reflect the goals of Phase 2 Codes, by protecting children from: • realistically simulated, generated and animated sexual content • high-impact text-based sexual content, including interactive services with a focus on chatbots and AI models which are built to provide pornographic content • nudity, making it clear that the protections apply to 'high impact' nudity only. eSafety also understands nudity classifiers are currently the primary tools for the proactive detection of pornography. Owing to this, **we consider applying the same measures to particular online pornography, online pornography and high-impact nudity is likely the most practical approach at this time.** eSafety is aware that a number of industry participants already use these measures to deal with this material in accordance with their own terms and conditions. We note there may be concerns about the potential for over blocking, if services are required to restrict access to pornographic material. We suggest services should take measures to minimise the potential for over blocking while also implementing adequate safety features. *At the time of publication there have been no specific or separate guidelines issued by the Classification Board to classify synthetic pornographic content generated by AI or animated pornographic content. In the absence of this guidance, eSafety considers industry may most

Type of material	Classifications	Suggested Considerations for proposed scope
		usefully proceed on the basis that it is possible such content could be class 1C or 2 material. A definition of 'high impact generative AI DIS' was adopted in the Phase 1 DIS Standard which industry may wish to consider adapting for use in the Phase 2 Codes.
Crime and violence and drug-related material	Class 2B – violence and themes of crime and drug and alcohol dependency (R18+ / Cat 1)	**Class 1B material relating to violence, crime, and drugs is covered by existing protections in the Phase 1 Codes and Standards.** This means that the worst of this content already subject to the Phase 1 requirements (which set out how that material should be dealt with in respect of all users, not just children). eSafety therefore proposes that additional steps to limit access to high-impact (but not class 1) crime, violence and drug-related material should be focussed on measures such as **having and enforcing policies about that material on services likely to be used by Australian children**.
Suicide, self-harm and eating disorder material	Class 2B – Themes of suicide, death, serious illness	We suggest an approach which provides appropriate protections and aligns as far as possible with protections applied in international jurisdictions**.** Due to the risk of harms associated with this material, we suggest **services which have a high or medium risk that they will be accessed by Australian child end-users** should **take reasonable steps to prevent and address harm** to children, and employ user-empowerment measures to allow all users to better protect themselves from this content. We suggest protections could include graduated, risk-based measures to address this material, such as filtering from children's feeds, downranking, content warnings, nudges, provision of support information, etc. We suggest these measures should also broadly align with requirements relating to this type of content which are now in place or will be introduced in the near future internationally (see further details in Table 1).
Simulated Gambling	Class 2B – as at Sept 2024, simulated gambling must be rated R18+ in the *Computer Game Guidelines*	We suggest a focus on material that has not already been through a classification process in accordance with the NCS,[81] particularly in circumstances where there is **a high risk that this content will be accessed by Australian children**. Online services should have and enforce policies about this material.

[81] If Notice Recipients choose to define a subset of services subject to minimum compliance measures in relation to simulated gambling or other material which may already have been subject to classification review processes, they could consider the Phase 1 DIS Standard's approach to 'classified DIS', which refers to classification by the Classification Board or according to an approved self-classification tool.

Type of material	Classifications	Suggested Considerations for proposed scope
Other content that may harm children	Class 2B – Other high-impact 'themes' as defined in the NCS (R18+ / Cat 1)	We suggest that all types of services could optionally have and enforce relevant policies, with only **services where there is a high risk that they will be accessed by Australian children** having a mandatory requirement to have and enforce such policies. eSafety acknowledges that several themes contemplated in the Classification Guidelines may be heavily dependent on their context to determine if they will be so high-impact that they are class 2 material. eSafety considers that industry should commit to implementing measures in their Codes which will improve safety tools to better account for context, to allow safety measures for children to be deployed more precisely. This proposed response to the remainder themes also recognises other pieces of legislation and schemes which already provide protections for end-users for material which relates to these themes (for example, anti-discrimination laws and criminal prohibitions on hate speech; and the eSafety Adult Cyber Abuse Scheme).

Overview of Childrens' access to Class 2 material online

The National Classification Code (which informs decisions made pursuant to each of the Classification Guidelines) states that classification decisions should give effect, as far as possible, to the principle that minors should be protected from material likely to harm or disturb them; and that adults should be able to read, hear, see and play what they want. Class 2 material through its designation in the Classification Guidelines, is therefore recognised as inherently harmful and disturbing to children, which is why it is age restricted. The Guidelines also note that this material may be offensive to sections of the adult community.

eSafety acknowledges the challenge of separating the potential impacts of online pornography, and other types of class 2 material from their broader societal context. For example, pornography can reflect and perpetuate gender inequality, sexism and violence, issues that already permeate many parts of our society. eSafety has previously considered this potential nexus in the Age Verification Roadmap.[82] Similar concerns apply to other types of class 2 material. In the present circumstances, the definitions and age restrictions for class 2 material as set out in the Classification Guidelines must be adhered to for implementing codes.

Given this legislated approach to class 2 material, we summarise and consider below evidence on how children access or encounter this material online and through which channels this occurs.

Pornography

References to 'online pornography' in this section and throughout this paper should be taken as a reference to 'particular online pornography' and 'online pornography' as set out in Table 2 above, as well as 'high impact nudity.'

Part II of the Age Verification Roadmap Background Report provides detailed evidence about children's experiences with online pornography, when and how they access it, and the potential impacts.

[82] eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023, p. 74.

Research indicates it is common for young people aged 16-18 to have viewed online pornography, both intentionally and unintentionally. The research also shows that almost half of children first encounter online pornography between the ages of 13-15 years old.[83]

The rates at which younger children encounter pornography are of particular concern to eSafety: in a 2022 survey, eSafety research found that 36% of boys and 40% of girls were first exposed to pornography before the age of 13;[84] and other studies report the average age for first viewing pornography is 13.6 years old.[85]

It is important to consider research insights into different levels of children's capacity for reasoning and decision-making at different ages. Stakeholders consulted for the Age Verification Roadmap agree that mainstream pornography poses a greater risk to younger children than to older teens.[86] Younger children often lack the ability to critically analyse pornography, and to temper its influence on their behaviours.

While eSafety research participants emphasised the importance of young people having agency over their own sexuality, they also highlighted that unexpected and unwanted encounters with online pornography are common. For example, 58% of surveyed young people who have encountered online pornography did so accidentally.[87] This includes when pornographic content is accidentally included in search results when a young person is searching for something unrelated.

Intentional access to pornography among young people is also significant. About 22% of young people surveyed first encountered pornography by searching for it online (30% of young men and 17% of young women).[88] For those whose first exposure was unintentional, the most common method was via internet pop-ups or Google search results.[89] Women were more likely to have their first exposure to pornography on social media than men (45% of young women and 30% of young men).[90]

Online pornography is not homogenous. The Age Verification Roadmap identifies a 'mainstream' form of pornography which targets a male heterosexual audience. This type forms the vast majority of online pornography accessed in Australia. Mainstream

[83] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 4.
[84] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 13.
[85] J Power, S Kauer, C Fisher, R Bellamy and A Bourne, *7th National survey of Australian secondary students and sexual health 2022*, The Australian Research Centre in Sex, Health and Society, La Trobe University, 2022.
[86] eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023, p. 47.
[87] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 23.
[88] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 15.
[89] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 15.
[90] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 15.

pornography websites, such as Pornhub.com and Xvideos.com – both referred to as 'tube' websites – consistently rank among the top 20 websites accessed in Australia. As of March 2024, these sites outranked popular platforms such as LinkedIn, OpenAI, and TikTok.[91]

Mainstream pornography tube sites are generally free to users and do not require registration. This increases their accessibility and risk profile compared to other types of online pornography that may be behind a paywall or require user registration.

However, mainstream pornography sites are not the only industry participants which must consider how they provide access to pornography for children. There is evidence that nearly every layer of the technology stack plays a role in how children may access or be exposed to online pornography.

70% of young people surveyed who accessed pornography did so on mainstream pornography websites. However, access via social media feeds (35%), ads on social media (28%), and social media messages (22%) is also common.[92] With increased device accessibility, children have more ways to access pornography. Visits to the mainstream pornography website Pornhub are most frequently conducted on smartphones, followed by desktop computers and tablets.[93] Pornographic sites are also frequently accessed on gaming consoles, such as Sony PlayStation.[94] Research published in February 2024 revealed that 39% of male participants reported daily viewing of pornography in the last 12 months, with 94% viewing it on an electronic device.[95]

eSafety considers effective measures to protect Australian children, especially younger ones, from accessing or being exposed to pornography online is integral to the aims of the Phase 2 Codes. These measures should be introduced across the technology stack, focusing on age assurance (set out in further detail in chapter 8).

Violence

Class 2 material includes high-impact violent content that lacks justification by context. The most extreme violent material is already addressed under Phase 1 Codes and Standards, which mandate its removal on report. Phase 2 Codes propose additional

---

[91] SimilarWeb, *Top Websites Ranking - Most Visited Websites in Australia (March 2024)*. It is notable that another web traffic analytics tool, SemRush, ranked Pornhub.com and xVideos.com at higher rankings than this, as the fifth and twelfth most accessed websites respectively in Australia in March 2024: see SemRush, *Most Visited Websites in Australia, Updated March 2024.*

[92] eSafety Commissioner, *Accidental, unsolicited and in your face - Young people's encounters with online pornography: a matter of platform responsibility, education and choice*, September 2023, p. 28.

[93] Pornhub Insights, *The 2022 Year in Review*, 2022.

[94] Pornhub Insights, *The 2022 Year in Review*, 2022.

[95] Crabbe et al, *Pornography exposure and access among young Australians: a cross-sectional study*, The Australian and New Zealand Journal of Public Health, 8 February 2024.

measures focussed on developing policies to protect children from class 2 violent content, augmenting existing protections from Phase 1.

A 2022 eSafety study found that more than a third (37%) of the Australian young people aged 14-17 who were surveyed said they had seen 'gory' or violent images or videos online, and one in five (23%) said they had seen violent sexual images or videos on websites or in online discussion.[96] While the research did not specify whether these images would classify as class 1 or class 2 material, it underscores the prevalence of children's exposure to violent content.

In March 2024, the UK online safety regulator, Ofcom, released a report titled *Understanding Pathways to Online Violent Content Among Children*.[97] Although conducted in the UK, the study examined children's interactions with global online platforms, including services that would be subject to any Australian Phase 2 code or standard. It provides insights into how Australian children might access or be exposed to such material online.

Ofcom's report highlighted several key findings regarding children's exposure to violent content online:

- **Ubiquitous exposure:** Children described encountering violent content as 'unavoidable'. Parents and carers expressed difficulty in keeping pace with the spread of this violent content. Professionals noted that parents often remain unaware of the content children encounter due to time constraints and varying levels of digital literacy.
- **Age of exposure:** Children reported first encountering violent content as early as primary school, with those aged 13-15 particularly noting the content as '*shocking*' or '*upsetting*'. Many children were encountering violent content without seeking it out, and were encountering it on '*most social media, video sharing and messaging services, as well as discussion forums, chat rooms and pornographic services*.' Many children admitted to creating user accounts on online services while under the minimum user age to access this content.
- **Accidental encounters:** Violent content was most commonly encountered accidentally through newsfeeds and recommendations, with children feeling they had no control over what was being recommended to them. Boys aged 13-15 were identified as the demographic most likely to actively seek out violent content online, often facilitated by recommender systems.

---

[96] eSafety Commissioner, *Mind the Gap - Parental awareness of children's exposure to risks online*, February 2022, p. 47.

[97] Ofcom, *Understanding Pathways to Online Violent Content Among Children*, March 2024; Ofcom, *Encountering violent online content starts at primary school*, March 2024.

- **Distribution channels:** Group chat distribution and inadvertent encounters during searches on video sharing platforms were also reported as common ways children came across violent gaming content.

Drug use

Research on online harms related to drug-related material is limited.

eSafety's research has found that more than a third (37%) of young people aged 14-17 surveyed said they had seen websites or online discussions where people talk about or show their experiences taking drugs.[98]

Page 169-178 of Ofcom's report *Protecting children from harms online - Volume 3: The causes and impacts of online harms to children* considers the harms to children when they are exposed to content about how to self-administer 'harmful substances.' This harm can manifest by encouraging children to use a particular substance, either through active or express encouragement.

Themes

The Classification Guidelines outline how specific 'themes' are categorised. These themes can include social issues such as crime, suicide, drug and alcohol dependency, death, serious illness, war, family breakdown and racism. Material related to these themes is prevalent across all sections of online industry, but its impact varies significantly depending on the context of its presentation. Recognising this variability, eSafety suggests industry adopt a graduated, risk-based approach for applying protections to these themes.

eSafety discusses in brief below two particular categories of 'themes':

- material relating to suicide, eating disorders and self-harm, given relevant international developments in addressing that material
- the theme of simulated gambling from the Computer Games Guidelines, given it is newly introduced compared to other themes and takes effect in September 2024.

Suicide, eating disorders, self-harm

Content related to suicide, eating disorders and self-harm has been identified as harmful to children in jurisdictions such as the UK and Ireland, as detailed in page 50-103 of Ofcom's report *Protecting children from harms online Volume 3: The causes and impacts of online harms to children*, and page 13-14 of Coimisiún na Meán's *Consultation Document: Online*

[98] eSafety Commissioner, *Mind the Gap - Parental Awareness of Children's Exposure to Risks Online*, February 2022, p. 46-7, 69.

Safety. Chapter 3 outlines how these jurisdictions are considering strict measures to limit children's access this such content.

eSafety has proposed specific measures for class 2 material that contains material relating to suicide, self-harm and serious illnesses such as eating disorders above at Table 2. These proposals aim to align with the emerging international standards addressing these issues, allowing industry to leverage technologies developed globally and to harmonise safety practices around the world. This also reflects the reality that many industry participants subject to Australian Phase 2 Codes will have obligations to restrict access to these types of material in these international jurisdictions, and may therefore have readily available tools and processes which can be leveraged to protect Australian end-users.

Suicide is specifically identified as a class 2 theme in the Australian Classification Guidelines, and therefore, we expect this to be explicitly addressed as outlined in our proposed scope for the treatment of Class 2 material. While eating disorders and self-harm are not expressly mentioned in the Guidelines, given the established evidence overseas regarding the harmful nature of content related to eating disorders and self-harm, and the likelihood that industry participants have or will develop tools to comply with international requirements, eSafety encourages industry to further consider during development of Phase 2 Codes whether material relating to eating disorders or self-harm could be categorised under the 'serious illness' classifiable element. This consideration would help extend protections to Australian end-users accordingly.

Simulated gambling in computer games

In September 2023, the classifiable element of 'Themes' within the Computer Games Guidelines was amended to include 'simulated gambling.'[99] From 22 September 2024, simulated gambling will be classified R18+.[100] This will mean it falls within the scope of Phase 2 Codes.[101]

Computer games are widely accessible by users of all ages. The IGEA 2023 *Australia Plays* report showed that 94% of Australian households have a device for gaming, and that 80% of children aged 1-17 play computer games.[102] Recent eSafety research also showed high rates of young people engaging specifically with online games.[103]

---

[99] eSafety notes that only simulated gambling is capable of being recognised as class 2 material under the current Computer Games Guidelines. Therefore, this paper does not consider any harms which may possibly arise from in-game purchases linked to elements of chance, as this is beyond eSafety's jurisdiction.
[100] *Guidelines for the Classification of Computer Games 2023 (Cth)*.
[101] *Guidelines for the Classification of Computer Games 2023* (Cth).
[102] Jeffrey E Brand et al., *Australia Plays 2023*, 2023, p. 4, 13.
[103] eSafety Commissioner, *Levelling up to stay safe: Young people's experiences navigating the joys and risks or online gaming*, February 2024, p. 16.

The *NSW Youth Gambling Study 2020*[104] surveyed children aged 12-17 and found they participated in simulated gambling through numerous channels, including simulated gambling apps, demo games, simulated gambling on social networking sites, or in-game betting with items, which correlates with increased likelihood of real-money gambling.

[104] Matthew Browne et al., *NSW Youth Gambling Study 2020*, January 2021.

8. Suggested model and measures for Phase 2 Industry Codes

Matters to be addressed by the Phase 2 Codes

As set out in chapter 5, on 1 July 2024, the Commissioner issued section 141 Notices detailing two matters in similar terms to the below for the online industry to address in producing Phase 2 Codes:

1. Protect and prevent children in Australia from accessing or being exposed[105] to class 1C and class 2 material.
2. Provide end-users in Australia with effective information, tools and options to limit access and exposure to class 1C and class 2 material.[106]

eSafety considers these matters are of substantial relevance to the community.

In determining these matters, eSafety has taken into account industry feedback and what it considers to be reasonable steps to deal with these issues. eSafety believes that all sections of online industry can address these matters and provide appropriate community safeguards through the respective Codes.

eSafety acknowledges that in reality, and despite all best efforts, access to the relevant material cannot in all cases be 'prevented'. Section 5.3 of the Phase 1 Codes head terms contains a provision that provided a degree of certainty for providers in this respect. eSafety expects that a similar provision should be usefully applied to future Phase 2 Codes.

eSafety does not expect every section of the online industry will have the same role in dealing with these matters. The actions required to provide appropriate community safeguards may differ across industry sectors and may overlap in some cases. However, eSafety expects that minimum compliance measures can and should be adopted by each section of industry to deal with each matter and create appropriate community safeguards as described in section 140(1)(d)(i) of the OSA.

---

[105] During consultations conducted in the production of the Age Verification Roadmap, eSafety received feedback that the word 'exposure' should be avoided where possible due to the potential to stigmatise online pornography as something inherently negative (see p. 77 of the Age Verification Roadmap). eSafety acknowledges this feedback but notes that the word 'exposure' is language already used in the *Online Safety Act 2021*, particularly in s 108(4)(a) in respect of Restricted Access Systems. To ensure continuity between the aims of the RAS and the Phase 2 Codes we have adopted the same language in the matters to be addressed in the Codes.

[106] The precise wording of each notice varies slightly to reflect the relevant industry section.

To that end, eSafety has set out suggested minimum compliance measures below. These measures should be considered by different sectors of industry or collectively by multiple sectors to deal with the matters specified in the section 141 Notices.

eSafety's preferred code model

As with the Phase 1 Codes, eSafety recommends that the Notice Recipients continue to adopt an outcomes-focused and risk-based approach when developing Phase 2 codes.

Outcomes focus

In the Phase 1 Codes process, desired outcomes were provided in the September 2021 position paper, published before issuing formal notices. The matters listed in the Phase 1 Notices were not directly linked to the outcomes in the position paper.

To streamline this process in Phase 2 and to emphasise the desired outcomes, eSafety has framed the matters in the current section 141 Notices as outcomes themselves. This retains the outcomes-focused approach while removing the need to formulate separate outcomes.

An outcomes-based approach provides industry participants with a common set of objectives and outcomes that align with the requirements of the OSA. It also offers the flexibility to implement measures that best suit their business models and technologies. These measures should be reasonable, proportionate and effective.

By phrasing the matters this way, every section of the online industry can take steps towards dealing with them. This addresses an issue from the Phase 1 Codes, where certain sections of the online industry determined that certain outcomes or objectives specified in the September 2021 Position Paper did not apply to them and therefore did not propose any measures to fulfil those objectives.

Given that the matters are now expressed so that all sections of the online industry can take steps to deal with them, Notice Recipients may consider options to develop fewer codes by combining some codes or aspects of codes. This is particularly relevant where the same measures can be adopted across different sections of industry (for example, account-based measures that could apply across interconnected ecosystems consisting of equipment, app distribution services, search engine services, SMS, RES, and DIS – see below).

Risk-based measures

Risk-based measures should be based on an assessment of both the risks an industry participant's services or devices present in exposing children to class 2 material, and the

unique roles they play across the tech stack, along with the opportunities they have to prevent and address access and exposure, including unwanted exposure.

This approach reflects relevant developments in international jurisdictions and was adopted by industry in Phase 1. For example, the Phase 1 SMS Code allocated services a tier rating based on risk factors such as the number of Australian end-users, total number of global active users, and discoverability functions. Minimum compliance measures within the Code were then limited to higher-risk services.[107] The risk-based approach was also carried over to the Phase 1 Standards, with both the RES and DIS Standards applying certain measures according to the risk profile of a service.

As specified in the Phase 1 position paper, eSafety expects that risk assessments for applying appropriate measures in industry codes should consider several relevant factors. For Phase 2, eSafety suggests that similar risk factors could be considered by industry. [108] Industry may also wish to consider the applicability to Phase 2 of risk factors applied in the Phase 1 Codes and Standards.

eSafety proposes that the application of minimum compliance measures in the Phase 2 Codes should consider the following risk factors:

- **The *purpose* of a service**, including:
    1. where the primary purpose of a service is providing class 2 material
    2. where the primary purpose of a service is not providing class 2 material, but it is permitted material
    3. where the service does not permit the provision of class 2 material.
- **The likelihood that an online service may be used to directly *expose* children to class 2 material.** This risk factor should reflect the likelihood class 2 material may be directly available to children on a service (such as SMS, RES, or DIS) without appropriate protections. This likelihood may not be analogous to whether class 2 material is permitted on a service.
- **The likelihood that a child will use a service to *access* class 2 material on a service**. This risk factor should account for the likelihood that online services such as equipment, search engines, app stores, hosting services or internet carriage services may be used by a child as a means of accessing class 2 material by providing connections to that material without appropriate protections.

[107] For example, see Measure 2 of the SMS code, which only applied to Tier and Tier 2 SMS.
[108] eSafety Commissioner, *Development of Industry Codes: Position Paper*, p. 50-51.

Measures that may be adopted in Phase 2 Codes to address the matters

1. Providing protections across every level of the technology stack

eSafety believes **all industry sections within the technology stack** can take steps to protect children from class 2 material and provide tools for adult users to limit their exposure to such content.

For some industry sections, this will include steps like confirming the age of users. eSafety discusses where age assurance measures could be applied below.

For other industry sections, the most appropriate steps will focus on implementing protective measures concerning class 2 material, building on reasonable and appropriate age checks.

We consider examples of some of the potential protective measures which industry may be able to adopt below:

Filters, **parental controls**, and **safety settings** could be applied at the equipment, internet carriage service and/or search level to prevent children from arriving at age inappropriate sites (ICSs can also promote third-party safety software, for example, through the Family Friendly Filter (**FFF**) scheme[109]).

There are differences between filtering, parental controls and safety settings.

Filtering can be applied by several sections of the technology stack as described above. Filters can:

- block entire websites containing categories of content (such as illegal, or age-inappropriate content) at the point-of-access
- be applied within individual online services, so the service is still accessible but certain material is subject to filtering.

Filters may have different effects. This can include removing certain content from view completely or restricting access and notifying the user that content has been filtered. They can introduce an interstitial notice about why the filtering has occurred, how to gain access to that content if they wish, or provide educational resources about the filtered content.

[109] See eSafety Commissioner, *Parental Controls*, 12 December 2024; Communications Alliance Ltd, *Family Friendly Filters*.

Filters and other automated solutions are not perfect, and carry risks of both under- and over-blocking. However, technologies' accuracy and error rates can be tested – as they are under the Family Friendly Filter scheme – and calibrated according to the circumstances. Stakeholders consulted for the Age Verification Roadmap spoke about the importance of measuring error rates and establishing review processes and solutions for those affected by potential errors. eSafety has also proposed industry adopt measures which are directed at measurably improving safety systems for end-users, including filtering systems.

Parental controls may include filters but encompass a broader range of measures, such as the capacity to manage or limit a child's online activity. eSafety is aware that many services currently provide options for parents and carers to alter default safety measures for the children in their household or care. Parental controls can also provide a more personalised approach than applying default safety measures for children. However, the responsibility for protecting children from class 2 materials should not fall solely on parents. The onus should not be on parents and carers to opt-in to safety settings for children. Instead, online service providers should implement child safety measures by default for children, with the option for parents and carers to change these defaults if they wish.

eSafety acknowledges that not all children have the benefit of a loving and supportive home life, and engaged parents and carers. eSafety considers this reinforces the need for industry to implement measures to protect and support children in relation to legally age-restricted material.

Safety settings encompass both filters and parental controls, along with additional measures. They can be applied by parents or guardians, or at the discretion of individual end-users. Safety settings may include limits on the types of activities a user can engage in and who can connect with them. Additionally, safety settings often include privacy controls, allowing users to decide what information about them is shared with others.[110]

The Age Verification Roadmap Background Report discussed various filtering, parental control and safety options available at the time of its publication. Since then, online service providers have introduced additional content filtering offerings.[111]

---

[110] See eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 260-270 for more detail on the measures outlined in this section.

[111] An examples of a newer filter system includes the equipment-level Sensitive Content Warning feature available on recently released Apple operating systems, which uses on-device machine learning to analyse photos and videos. Apple states that 'if Sensitive Content Warning determines [a user has] received a photo or video that appears to contain nudity, it blurs the image, displays a warning that the content appears to be sensitive and offers ways to get help.' Users who have declared that they are adults while setting up their Apple devices (through the process of creating an Apple ID) may opt in to this feature: support.apple.com/en-au/105071. A similar feature called

Other measures

Hosting services and app distribution services could require the services they host to **comply with age confirmation and other requirements**, and can **cease hosting/providing access** if they receive a complaint that cannot be resolved with the service.[112]

Hosting services provide data storage and other computing support to enable sites to exist on the internet. These services can set terms of service for the sites they host. eSafety suggests these terms could include requirements for customers to comply with the local laws in jurisdictions where hosted websites are being accessed. For example, hosting services may require website owners to take reasonable steps to prevent children's access to online pornography and to comply with relevant Phase 2 Codes. If a site fails to comply, the hosting service can provide the site with an opportunity to rectify the issue before considering enforcement options such as suspending its services.[113]

For example, Amazon Web Services, Google Cloud and Microsoft Azure have policies in terms which allow them to suspend or terminate access to websites at any time, including if usage violates applicable laws and regulations.[114]

App distribution services, such as the Apple App Store and Google Play, allow users to download applications onto their devices. Both platforms use a review process where they assign age-based content ratings and enforce terms of service that prohibit apps whose main purpose is to provide pornography.[115] According to their policies, apps that violate these terms are rejected and not made available for download. It's important to note that the Apple App Store and Google Play Store require users to log in with centralised accounts (Apple ID or Google account) enabling them to manage app purchases and settings across devices.

Search engines could apply safe search functions by default, and downrank or demote online services that do not comply with regulatory requirements.

'Safe search' functions automatically remove content such as pornography and graphic violence from search results returned by search engines. These functions are already

---

Communication Safety is also available for child end-users who have declared that they are children, and who are associated with an Apple Family Sharing Group. Communication Safety uses the on-device machine learning to analysis messages which are both sent and received for any nudity: support.apple.com/en-us/105069.

[112] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 299.

[113] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 299.

[114] Amazon Web Services, *AWS Service Terms*, 7 May 2024, 1.4; Amazon Web Services, *AWS Acceptable Use Policy*; Google Cloud, *Google Cloud Platform Terms of Service*, 4; Google Cloud, *Google Cloud Platform Acceptable Use Policy*; Microsoft, *Acceptable Use Policy*, February 2011.

[115] Apple, *App Store Review Guidelines*, 1.1.4, 1.2; Google, *Inappropriate content*, Play Console Help.

available through major search engines such as Google and Bing.[116] It is also possible to enable these settings by default or based on system information; for example, currently, Google applies SafeSearch settings by default for children who are signed into an account managed with Family Link and those who are under 13.[117]

Downranking is already commonplace among many search engines in response to violation of their policies.[118] For example, Google explicitly states it will detect (through automation and human review) violations of its spam policies which can result in search result demotion, or complete removal from search results. Possible spam policy violations include:[119]

- purchasing and re-registering expired domains to manipulate search rankings
- hosting hacked content on a website
- using hidden text or links that aim to manipulate search engines
- engaging in link spam, which can determine search relevancy
- hosting malware or enabling malicious behaviours on a website.

Google's approach to removing content from search results encompasses various behaviours that can lead to 'demotion or removal', including scams, fraud, personal information violations, and legal claims such as copyright law, defamation, counterfeit goods, and other removals which may be ordered by a court.

eSafety suggests that Google's approach to removing copyright materials could serve as a model for addressing how class 2 material appears in search results. When Google receives a 'high volume' of valid requests to remove copyrighted material, it uses this data to lower the ranking of infringing material in-search results. This approach aims to reduce users' exposure to copyrighted material that violates laws, promoting the visibility of original, non-infringing material instead.

A comparable reporting-and-demotion approach could be applied for end-users to report websites that fail to comply with Industry Codes. For example, if end-users are given the means to report websites that host pornographic material without implementing required safety measures, these websites could be penalised in search results. By demoting non-compliant sites, users would be more likely to encounter compliant websites with relevant protections through search engines.

---

[116] Google, *Google SafeSearch;* Microsoft, *Turn Bing SafeSearch on or off*.
[117] Google, *Google SafeSearch*.
[118] In addition to Google, the search engine Bing also sets out numerous factors which it uses to determine whether a website should be downranked.
[119] Google Search Central, *Spam Policies for Google web search*, 5 June 2024.

eSafety notes that the Phase 1 Search Code includes a provision that requires search engine services to implement processes that limit exposure to class 1B material in search results. Guidance for this provision suggests services design ranking algorithms which mark certain regulated content as 'less authoritative' and less likely to appear in search results or summaries. The Code also requires search engine services establish mechanisms to handle reports from Australian end-users regarding illegal class 1A and 1B materials. Guidance for this provision emphasises the need for these mechanisms to facilitate appropriate actions based on factors such as the urgency and extent of harm, the effectiveness of different interventions, and the source of the reports. eSafety expects that similar considerations would apply to search engines handling demotion requests for non-compliant sites under Phase 2.

Case studies of protections across the technology stack

To illustrate how protections across the technology stack can reinforce each other and strengthen overall safety measures, eSafety has included case studies in Annexure A of this position paper. These case studies demonstrate real-world scenarios where protections against class 2 material could be effectively implemented for end-users. The case studies are illustrative only. They are not intended to indicate eSafety's compliance and enforcement approach in relation to a code or codes.

2. Leveraging digital ecosystems for privacy-protecting, data-minimising age assurance and complementary safety measures

eSafety believes that existing internet ecosystems can effectively leverage existing end-user information gathering processes for privacy-protecting, data-minimising age assurance and complementary safety measures. This approach aligns with safety-by-design guidance, including to 'leverage the use of technical features to mitigate risks and harms.'[120]

'Ecosystems' of online products and services can exist where they are interconnected through:

- access, integration and interoperability
- defaults and pre-installation
- user accounts
- bundling and typing strategies.[121]

[120] eSafety Commissioner, *Safety by Design*, April 2024.
[121] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 94.

The ACCC's *Digital Platform Services Inquiry Interim Report No. 7 – Expanding digital platform ecosystems* (September 2023) explores the prevalence and operation of digital ecosystems. It identifies how digital platform service providers interconnect their products and services, noting both positive and negative implications for consumers.[122] The report highlights that online service providers already collect and share user data across their various services and products.[123] These existing processes and connections could be leveraged to support online safety, including preventing children from encountering class 2 material.

Most relevant to this position paper, the ACCC considers scenarios where an end-user maintains a centralised 'user account' to 'enable interconnections between products and services in a number of ways.' User accounts streamline consumer access to products or services within an ecosystem, eliminating the need for separate accounts for each service. They allow 'single sign on' across multiple products or services, both within and beyond a specific ecosystem. They may also allow consumers to synchronise account usage and preferences across devices.[124]

In the context of the OSA, these ecosystems may consist of products and services that may span across many sections of the online industry. Despite spanning multiple sections, a centralised user account consolidates user information in one place, including passwords, email addresses, profile pictures, phone contact details, date of birth, payment details and security details.

Centralised user accounts are already used to gather data across online services for targeted advertising purposes.[125] The ACCC also notes that collection of end-user account data (which regularly includes age information) enhances the delivery of relevant recommendations and search results. [126] This data collection is already enabled and protected through the terms of service of those ecosystems, extending across interconnected systems and products.[127] Companies such as Amazon, Apple, Google, Meta

[122] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, chapter 5.2.
[123] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, chapter 7.2.1
[124] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 96.
[125] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 151. eSafety also notes that in the Government has agreed-in-principle and noted several proposals in the *Privacy Act Review Report* which may qualify the extent to which user data is collected by online service providers: see p 32-33.
[126] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, *September 2023, p. 96.*
[127] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 157-158.

and Microsoft obtain user data through features such as linked or centralised user accounts for use elsewhere in their ecosystems.[128]

eSafety does not support invasive, unreasonable or unfair data collection practices. It is important to clarify that nothing in this position paper regarding account-level age assurance and complementary safety measures flowing through product and service ecosystems should be taken as eSafety's endorsement of additional end-user data collection by online service providers beyond what is essential for their operations. Instead, eSafety advocates using existing systems that currently generate revenue for online service providers to also deliver effective and interoperable safety mechanisms to end-users where possible. This may include implementing appropriate and reasonable age assurance measures, such as those proposed for inclusion in the Phase 2 Codes.

So long as any data collected is used in a lawful, reasonable and fair way, eSafety believes introducing age assurance measures does not significantly change the current data collection relationship between users and ecosystems. In fact, effective age assurance may reduce the volume of data collected from child end-users, particularly where legal requirements in different jurisdictions restrict data collection from children and mandate high privacy settings by default for users accurately identified as children, as discussed in relation to the UK Age Appropriate Design Code below.

Users may need to log in to their centralised user account to access the full functionality of certain equipment. This could involve logging in during device setup, updating software, or downloading new apps. Tying age-based safety (and privacy) measures to these accounts and their ecosystems can reduce the need to apply new settings each time a user switches devices or moves across services.

For example, end-users' ages could be checked and safety features could be enabled by default on their centralised user account that extends to all interconnected services within the ecosystem or accessed through that account.[129] Users may also create 'family accounts' within an ecosystem, allowing adults to either opt in or opt out of parental controls on devices used by children in their care, provided these devices are part of the same ecosystem as the adult's devices.

This approach could streamline age assurance across multiple platforms and services, reducing the need for users to confirm their age repeatedly. Where practicable, these settings could migrate with the user across different services, facilitating the effective use of other safety tools, such as age-tailored filters, settings and accounts.[130] Downstream

[128] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 96.
[129] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, chapter 6.
[130] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 286.

services could potentially leverage upstream age checks using this approach. For example, app distribution services could use age checks conducted on the centralised user account to prevent the download of age-inappropriate apps, and interconnected search engines could automatically apply safe search settings.

One way to achieve this through the Phase 2 Codes could be to identify a set of core provisions common to all or some of the codes, guided by the position paper Minimum Compliance Measures. The approach could emphasise account-level settings for large service providers whose products and services span:

- manufacturing and supplying equipment
- providing an app distribution service and/or a search engine service
- providing an SMS, RES, or DIS.

The ACCC notes that ecosystems may exacerbate consumer 'lock-in'[131] and that children are particularly vulnerable to this as they may be exposed to an ecosystem at a young age, for example through the device they are provided with at school.[132] Accordingly, any measures implemented should allow child end-users to update settings in a way that promotes switching between ecosystems when they are no longer underage. eSafety also suggests in chapter 9 that ecosystem-based age assurance measures should be interoperable with other online services to the extent possible, further reducing consumer lock-in.

Privacy preserving measures

eSafety recognises concerns that age assurance technologies and associated protective measures contemplated for the Phase 2 Codes may impact users' privacy. eSafety acknowledges these risks and encourages industry to consider how to adopt measures in Phase 2 that also preserve end-user privacy. The Age Assurance Trial will specifically consider privacy and security as they relate to different age assurance measures.

eSafety believes age assurance can enhance privacy, especially for children. Some safety and privacy measures are only effective when the person's age is known, as specific privacy-preserving or safety measures may apply only to children. [133] For example, if a child claims to be an adult, they may be exposed to age-inappropriate data collection and advertising practices.[134]

---

[131] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 168.
[132] ACCC, *Digital Platform Services Inquiry – Interim Report 7: Report on expanding ecosystems of digital platform service providers*, September 2023, p. 155.
[133] ACCC, *Digital Platform Services Inquiry – Interim Report No. 8: Data products and services*, March 2024, p. 109.
[134] ACCC, *Digital Platform Services Inquiry - Interim Report No. 6: Social media services in Australia*, March 2023, p. 154.

International jurisdictions have recognised the interconnectedness between age assurance and effective privacy measures. For example, Standard 3 of the UK Age Appropriate Design Code sets out that online services should 'establish age with a level of certainty that is appropriate to the risks to the rights and freedoms of children that arise from your data processing, or apply the standards in this code to all your users instead'.[135] By establishing the age of users with a level of certainty, online services can better tailor protections and safeguards for users' personal data (which, under the UK Code, includes defaulting children's data settings to the highest possible privacy settings). The UK Information Commissioner's office acknowledges 'there is a tension between age assurance and compliance with [privacy obligations] as the implementation of age assurance could increase the risk of intrusive data collection…'. However, it also notes that age assurance and compliance with privacy obligations are compatible if privacy by design solutions are used.[136]

As noted in chapter 6, in response to the Australian Privacy Act review, the Australian Government has agreed that a Children's Online Privacy Code should be developed. This code could align with the UK's Age Appropriate Design Code.[137]

eSafety has also proposed in chapter 9 that industry could adopt privacy impact assessments as part of implementing age assurance measures, which should include consideration of privacy-by-design principles.

3. Building on pre-existing regulatory schemes which aim to protect and prevent children from accessing class 2 material

In consultations with industry prior to publishing this Phase 2 position paper, eSafety found that industry prefers flexibility in adopting measures across different industry sections to achieve the two matters.

eSafety believes industry should consider pre-existing protective measures within the OSA related to class 2 material, the Restricted Access System Declaration (**RAS Declaration**), and the BOSE Determination. These schemes are designed to be flexible while maintaining strong protection expectations for the Australian community. These regulatory schemes are discussed in chapter 6.

eSafety also seeks to build on the work of the Phase 1 Codes and Standards. Importantly, this includes recognising where different industry sections have varying capacities to fulfil the matters set out in the section 141 Notices as compared to measures they adopted for

[135] *Information Commissioner's Office*, 3. Age Appropriate Application, September 2020.
[136] *Information Commissioner's Office*, 3. Age Appropriate Application, September 2020.
[137] Australian Government, *Government Response – Privacy Act Review*, September 2023, p. 30.
[138] *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* Explanatory Statement.

Phase 1. This also acknowledges the broader scope of content and different harms addressed by Phase 2.

The RAS Declaration

The RAS Declaration has four basic components:

- Ask end-users for their age
- Take reasonable steps to confirm age
- Deny access to class 2 material for those under 18
- Provide safety information

To prevent any inconsistency between what a service provider might be required to do under the RAS Declaration and an industry code or standard, eSafety believes the RAS should set the baseline requirements for SMS, RES and DIS. Any further measures established through industry codes or standards should build on and remain consistent with the RAS's four key components.

The Explanatory Memorandum to the *Online Safety Bill 2021* sets out that:

> *The purpose of a restricted access system declaration is not to prevent access to age-restricted content (whether it is user-generated content or otherwise) via any platform, but to seek to ensure that:*
>
> - *access is limited to persons 18 years and over in the case of R 18+ content; and*
> - *that the methods used for limiting this access meet a minimum standard*.

The aim of the RAS is to **prevent** children accessing content rated R18+ or higher (i.e., class 2 material). The Memorandum also states that the Commissioner must consider the objective of '**protecting** children from exposure to material that is unsuitable to children' when making an instrument under the RAS provisions of the OSA.

This language of **prevention** and **protection** is consistent with the first Matter to be addressed by the Notice Recipients in the section 141 Notices.

The BOSE Determination

The BOSE Determination sets out expectations for all social media services, relevant electronic services, and designated internet services. Section 12 emphasises the core expectation that providers will take reasonable steps to implement measures – technological or otherwise – to prevent children accessing class 2 material. It specifically mentions that reasonable steps to comply with section 12 could include 'implementing age assurance mechanisms.' The *Online Safety (Basic Online Safety Expectations) Amendment*

Determination 2024 (the **2024 Amendment Determination**) updated this provision to refer to '*appropriate*' age assurance mechanisms.

The Explanatory Statement to the 2024 Amendment Determination clarifies that the term 'appropriateness' in the context of age assurance measures should take into account these factors:

- the effectiveness of the age assurance mechanisms
- the extent to which class 2 material is provided on the service
- the likelihood of children accessing the material on the service.

The Explanatory Statement to the 2024 Amendment Determination[138] says this means in some instances, asking users to self-declare their age or date of birth may provide an effective signal or barrier to unintentional access by children. In other instances, services will be expected to establish a user's age with a greater certainty, appropriate for the level of risk of the material that can be accessed on the service.

The factors relevant to determining appropriateness under the BOSE Determination could also be adapted into any risk-tiering system adopted in the drafting of the Codes. Further, to ensure consistency between the RAS and BOSE schemes, eSafety's Regulatory Guidance on the BOSE recognises that elements of the RAS Declaration could be adopted in taking reasonable steps to comply with section 12 where there is likely to be class 2 material.

Building on the RAS and the BOSE

Aligning Phase 2 Codes with existing protective schemes such as the RAS and the BOSE offers several advantages. It provides industry with a familiar framewor, leveraging regulatory guidance and resources already in place to support these measures. This alignment can facilitate smoother implementation and response to new Codes that broadly mirror the principles of the RAS and the BOSE.

During informal consultations with eSafety prior to this position paper's publication, industry participants voiced concerns about the complexity and regulatory burden of complying with different schemes such as the RAS and the BOSE in addition to industry codes. To address these concerns, eSafety believes industry, as the leaders of the drafting process for Phase 2 Codes, should consider adopting measures that align with existing schemes, where feasible. This approach not only enhances clarity and reduces compliance challenges, but it also promotes more uniform adherence to regulatory requirements across different schemes, such as aligning Phase 2 Codes requirements with the BOSE.

[138] *Online Safety (Basic Online Safety Expectations) Amendment Determination 2024* Explanatory Statement.

eSafety acknowledges that existing schemes like the RAS Declaration and the BOSE may not perfectly align with the comprehensive aims of the Phase 2 Codes. For example, the RAS currently applies only to R18+ material on SMS, RES and DIS provided from Australia, whereas the Phase 2 Codes aim to cover a broader range of material and services accessed by end-users in Australia. Nevertheless, by building on the foundation of established compliance schemes, industry can advance protective measures while adapting them to suit the broader scope envisioned by the Phase 2 Codes.

eSafety also emphasises the need for a variety of complementary safety measures beyond age assurance. These suggestions are set out in chapter 9 and Annexure A.

The Phase 1 Codes

The Phase 1 Codes address certain materials that are similar to those within the scope of the Phase 2 Codes. They establish minimum compliance measures that may overlap in some circumstances with the content considered under Phase 2.

With the ecosystem approach in mind, eSafety identifies distinct roles within sections of the online industry for both Phase 1 and Phase 2 Codes. eSafety outlines these differences below and explains why these differences are appropriate based on the specific subject matter and harms targeted by Phase 2.

App distribution platforms

Measure 3 of the Phase 1 App Distribution Platform Code requires app distribution platforms to 'make age and/or content ratings information about third-party apps available on the app distribution service to Australian end-users at the time those third-party apps are released on the app distribution service.'

The Age Verification Roadmap and Background Report found that app distribution services serve as gatekeepers for app access and can provide meaningful safeguards around the types of apps children can access, including social media and games. However, current practices show that app distribution platforms do not effectively enforce age ratings. Although they collect user age information based on self-declaration during account creation, they do not restrict access to apps by default based on these ratings. For example, users aged 13 to 16 can download apps rated 17+ on the Apple App Store unless specific parental controls are activated.[139] As a result, age ratings on app stores currently act as a guide only.[140]

[139] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 291.
[140] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 99.

Due to the role app distribution platforms play in controlling children's access to sources of class 2 material, eSafety proposes measures for Phase 2 that aim to ensure that age rating systems be enforced on app distribution platforms, rather than solely notified to end-users. This could include steps to confirm end-users' ages and applying default restrictions to their ability to download apps.

Equipment

The equipment section of the online industry pertains to devices used by end-users in Australia for SMS, RES, DIS and any internet carriage service. This includes manufacturers, suppliers, and those who maintain and install equipment used to access online services.[141]

Under the Phase 1 Industry Codes, equipment services are categorised into three risk tiers. Tier 1 encompasses equipment posing the greatest safety risks relevant to class 1A and class 1B material. Equipment classified as Tier 1 interactive children's devices are those 'primarily intended for use by, primarily designed for, and/or primarily marketed to, children. They are subject to additional measures to enhance protection, such as defaulting to the highest possible location and privacy settings on operating systems.

eSafety does not consider that drawing a distinction between equipment targeted at children and equipment that may be used by children is useful for the purpose of Phase 2 measures. This is based on user traffic information published by Pornhub, Australia's most visited pornography website. Pornhub's data shows that in 2023, 86% of Australian traffic to the website came from phones. The site was also accessed through various web browsers on both desktop and on mobile platforms, as well as from game consoles.[142]

None of these devices are typically marketed specifically to children, highlighting the need for broader protections under Phase 2 Codes. Moreover, many younger children use shared family devices rather than having their own individual devices, as highlighted in the Age Verification Roadmap. Therefore, focusing on child-targeted devices alone does not adequately address the practical reality of device use among children in relation to class 2 material such as pornography.

eSafety proposes that the Phase 2 Equipment Code should build on the 'Tools' described in Measure 6 of the Phase 1 Code. Industry should adopt measures that not only allow Australian end-users to mitigate risks to children, such as allowing the creation of parent and child accounts or requiring parental activation of safety controls, but also by requiring device manufacturers to implement default measures that reduce the risk of harm. These

[141] *Online Safety Act 2021* (Cth), s 135(h). Within the definitions of the Codes, equipment includes actual devices as well as operating systems which are loaded onto devices.
[142] Pornhub Insights, *2023 Year in Review*, 2023.

measures could include default device-level safety settings and filtering of inappropriate content, which could be supported by device-level age assurance mechanisms.

This approach is supported by evidence from the Age Verification Roadmap for two main reasons.

First, research shows that relying solely on parents or carers to opt-in to device-level protections is ineffective. In Australia, 45% of parents and carers do not use parental controls due to reasons such as generational gaps in digital literacy, cost concerns, and inability to calibrate settings.[143] International research shows high adoption rates of optional on-device safety tools by parents, highlighting their importance.[144] eSafety encourages the online industry to build on its efforts to comply with Measure 5 and optional Measure 8 in the Phase 1 Equipment Code, which provides information about safe use of equipment online so that this may increase in Australia.

Second, device-level measures are likely to have added benefits due to their flow-on effects through a device ecosystem. For example, in the Age Verification Roadmap, stakeholders emphasised that options within device and operating system are the most scalable, straightforward and comprehensive for age assurance. They streamline settings and permissions across various sites, supporting privacy by reducing the need for data collection across multiple services. Stakeholders also noted that consumers are accustomed to providing personal information, including age, to companies when buying and setting up devices, especially those devices which are associated with a technological ecosystem. This familiarity may facilitate easy transfer of information across multiple online services.[145]

Hosting services

Industry codes or standards developed under the OSA apply to services that host material in Australia.[146] The Phase 1 Codes, developed by industry, distinguish between two categories of hosting services: first-party and third-party.[147] The Phase 1 Hosting Services Code applies only to providers of third-party hosting services. First-party hosting services that also offer social media services, relevant electronic services or designated internet services are regulated by Phase 1 Industry Codes or Standards for those specific services.

eSafety generally agrees that the Hosting Code should continue to focus on third-party hosting services. The Phase 1 Code helpfully recognises that these services do not have a direct relationship with end-users, but they do have a direct relationship with their own

[143] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 266.
[144] Ofcom, *Children and Parents: Media Use and Attitudes Report*, 19 April 2024, p. 49.
[145] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 171.
[146] *Online Safety Act 2021* (Cth), s 134(f).
[147] Hosting Services Online Safety Code (Class 1A and Class 1B Material), s 2(b).

customers. These customers may be the providers of the SMS, RES, and DIS accessed by Australian end-users.

eSafety encourages industry to include in any Phase 2 Hosting Services code measures for establishing and enforcing policies or terms of service. These should require their customers to comply with local laws and enforceable industry codes or standards in jurisdictions where their material is hosted. This obligation is already commonly applied among many large internet hosting services, and aligns with Measures 1 and 2 of the Phase 1 Hosting Services Code, outlined earlier in this chapter.

Search engines

Most visits to websites start with a search engine.[148] The Age Verification Roadmap found that search engines, like app stores, can act as gatekeepers to reduce children's access to age-inappropriate materials via search.[149]

Stakeholders in the Age Verification Roadmap consultation raised concerns about how search results can deter users from accessing compliant websites. When a website implements safety measures as required by an enforceable code or standard, children might simply migrate to non-compliant websites that are easier to access. Search engines may then prioritise these non-compliant sites in their results, reflecting the change in traffic.

eSafety hopes that adopting the account-based measures that could apply across interconnected ecosystems suggested in this position paper will help industry reduce the risk of promoting non-compliant websites by introducing barriers to children's access at multiple points within the ecosystem. For example, if a device is linked to an end-user who has not completed age assurance for their centralised user account, images of class 2 material could be blurred by default, or websites tagged with RTA tags or known to contain class 2 material could be automatically filtered out of search return results. To complement these features, search engines may consider adopting measures in the Phase 2 Codes that make their safety tools (already required under Measure 9 of the Phase 1 Internet Search Engine Code) the default setting for end-users who have not completed age assurance measures to confirm they are an adult.

eSafety also suggests building on measures for de-listing websites based on user reports, as outlined in Measure 10 of the Phase 1 Search Code. eSafety believes search engine services should allow users to report websites providing class 2 material that do not comply with requirements established under the Phase 2 Codes or the RAS Declaration.

---

[148] Vinny La Barbera, *8 SEO stats that are hard to ignore*, imFORZA.
[149] eSafety Commissioner, *Age Verification Roadmap: Background Report*, August 2023, p. 94.

This could allow search engines to ensure these non-compliant websites are not prioritised over compliant sites in search results.

The Phase 1 Standards

The Phase 1 RES and DIS Standards contain requirements which eSafety encourages industry to align with, as relevant and as possible, in any Phase 2 Codes.

eSafety also considers that certain requirements in the Phase 1 RES and DIS Standards reinforce the goals of the Phase 2 Codes.

For example, the DIS Standard sets out requirements for online services assessed as Tier 1 under the Standard. Tier 1 services include 'high impact DIS' which have the sole or predominant purpose of enabling access to high impact materials, or make available high impact material that has been posted by users. Tier 1 services must:

- make clear in terms of use that an Australian child is not permitted to hold an account on the service
- take appropriate action to ensure that a child in Australia known by the provider to be under the age of 18 does not become an end-user of the service
- take appropriate action to stop access by a child in Australia known by the provider to be under the age of 18.

These requirements are similar to protective measures suggested for Phase 2, such as age-gating websites wholly or partially dedicated to providing class 2 material (which is high impact by definition). Additionally, proposed measures relating to age assurance in Phase 2 may enable more accurate application of these Phase 1 Standards requirements which apply specifically to children.

9. Suggested measures for the Phase 2 Codes

Considering the points in chapter 5, eSafety has compiled a list of suggested measures for different sections of the online industry. eSafety has also included user case studies in Annexure A which contain examples of measures that industry could adopt.

While this list is more comprehensive than the measures suggested in the September 2021 position paper for Phase 1, it is not exhaustive. eSafety encourages industry to consider any additional measures that could help achieve the matters in the Phase 2 Notices.

Further, eSafety does not propose to set out suggested measures other than those which specifically relate to class 2 material. However, the industry codes can and should include supportive measures for dealing with class 2 material, such as:

- establishing mechanisms for eSafety to require reporting about compliance and end-user uptake and effectiveness of measures implemented under the Codes
- promoting knowledge-sharing among industry participants, including mandatory industry forums
- facilitating knowledge-sharing with civil society groups, public interest groups, and representatives of marginalised communities to ensure equitable application of Codes measures
- providing end-users with information about safety features, educational resources, and links to complaint systems (both those administered by industry participants and by eSafety)
- completing risk assessments for new technological features in online services that could affect end-users' exposure to class 2 material, with requirements for consultation with eSafety about these features.

As industry has already demonstrated its ability to establish effective supportive measures in the Phase 1 Codes, eSafety does not believe industry needs further guidance on these measures in this paper.

eSafety considers that the industry-led, co-regulatory Industry Codes scheme in the OSA allows industry bodies to apply their expertise and technical understanding to develop robust codes. This belief is supported by the Phase 1 process, where six industry-developed codes were successfully registered. Pursuant to the expectations in the OSA, eSafety consulted with Notice Recipients throughout the two-year development of Phase 1 Codes

from September 2021 to September 2023. This was followed by more than six months of specific consultation on the Phase 2 Codes starting from November 2023.

This ongoing consultation, along with the work already done by industry in Phase 1, will provide a solid foundation for developing Phase 2 Codes. Building on the foundational work – such as establishing a robust set of head terms and definitions in Phase 1, eSafety is confident that Notice Recipients are well-equipped to develop Phase 2 Codes within the timeframes set out in the section 141 Notices.

One important part of the Notice Recipients' role in drafting the Phase 2 Codes will be to consult with industry participants and with the public. eSafety recognises that certain sections of the Australian community face higher online risks than others, and that these harms – and measures intended to address these harms – can have differential impacts, especially for disadvantaged or marginalised groups with intersecting risk factors. This includes, but is not limited to, First Nations people, culturally and linguistically diverse communities, people with disability, people who identify as LGBTIQ+, as well as, depending on the circumstances, women, older people, and children and young people.

Children and young people, in particular, are at greater risk due to the social, emotional, psychological, and sometimes physical impacts that can result from encountering harmful content and behaviour online, especially when such content is encountered unintentionally. With these disproportionate impacts in mind, the public consultation led by the Notice Recipients should aim to be inclusive, seeking feedback across various disciplines and a wide spectrum of lived experiences.

eSafety expects that the Phase 2 Codes and any resulting measures will consider these important issues, while also ensuring that proposed requirements do not unduly restrict communities and individuals in their online expression or make them feel marginalised. Research has also shown that designing protective measures that centre and account for the experiences of marginalised communities can lead to better outcomes for all users.[150] eSafety has already considered the needs of diverse communities concerning material covered by the Phase 2 Codes in the Age Verification Roadmap and Background Report. eSafety encourages industry to carefully consider the insights from these documents when developing the Phase 2 Codes.[151]

---

[150] Afsaneh Rigot, *Design From the Margins: Centering the most marginalized and impacted in design processes - from ideation to production*, 13 May 2022.
[151] eSafety Commissioner, *Roadmap for Age Verification*, March 2023, p. 65; eSafety Commissioner, *Roadmap for Age Verification: Background Report*, August 2023, p. 39, 127-128, 275, 278, 316-317.

Suggested minimum compliance measures which should be considered for the Phase 2 Codes

The examples of suggested measures set out below to address the matters contained in the Notices are not exhaustive. Industry is encouraged to have regard to the issues discussed in Chapter 8 above. Industry should also consider adopting any additional measures it considers appropriate to address the Matters.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
1.0	**Online services must take reasonable steps to ascertain the age of users by implementing appropriate age assurance measures before providing access to class 1c and class 2 material.**[152]

1.1	RES DIS SMS EQP SES APP
	Proportionality/risk
	The appropriateness of age assurance measures applied should be proportionate to risk. Risk factors may include, for example:
	• the **purpose** of a service, including:
	  ○ where the primary purpose of a service is providing class 2 material
	  ○ where the primary purpose of a service is not providing class 2 material, but it is permitted material
	  ○ where the service does not permit the provision of class 2 material.
	• **The likelihood that an online service may be used to directly *expose* children to class 2 material.** This risk factor should reflect the likelihood class 2 material may be directly available to children on a service (such as SMS, RES, or DIS) without appropriate protections. This likelihood may not be analogous to whether class 2 material is permitted on a service.
	• **The likelihood that a child will use a service to *access* class 2 material on a service**. This risk factor should account for the likelihood that online services like equipment, search engines, app stores, hosting services or internet carriage services may be used by a child as a means of accessing class 2 material by providing connections to that material without appropriate protections.
	
	Terms and conditions
	Terms and conditions of use for online services should require users to comply with age assurance measures before being able to access class 2 material on that service.
	
	Online services should have, and enforce, clear actions, policies or terms of service relating to class 2 material. Terms of service do not need to prohibit this material in its entirety, but

[152] Numerous forms of age assurance technologies are explored in the eSafety *Age Assurance Issues Paper* and *Age Verification Roadmap: Background Report*, see Chapter 8.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
	where it is permitted, terms of service should include clear direction on how child end-users of a service will be protected and prevented from accessing or being exposed to this material. eSafety suggests that policies and actions should be implemented in respect of different kinds of class 2 material by adopting the graduated, risk-based approach set out in **Table 1**.
2.0	**Online services should ensure age assurance measures are user-friendly, privacy-preserving and easy to adopt.**
2.1	RES DIS SMS EQP SES APP **Ecosystem approach** Age assurance measures could be applied through centralised user accounts in 'ecosystems' of online products and services so that: • age assurance can be conducted in a data-minimising, privacy-preserving manner by being layered into systems where users already provide their personal information, commonly including age information; • age-related settings can be carried across multiple aspects of any online ecosystem; • users may be as minimally inconvenienced by age assurance measures as possible by occurring at a single point, to reduce compliance and 'consent fatigue'[153]; • the increased technical capabilities of large, pre-existing ecosystems of online services can be applied to age assurance systems to benefit end-users including: o better developed security measures o more developed safety settings o better information privacy options for users **Examples of points where age assurance measures may be applied (non-exhaustive)** • When a user creates a centralised user account which carries across an 'ecosystem' of interrelated products and services. • Where users are asked to log in and provide personal details for the purpose of connecting a centralised user account to a piece of equipment, such as a device or operating system. • Where users are asked to provide centralised account details for the purpose of accessing the full functionality of an online service, for example: o where activating a new device for the first time, and attempting to access the full functionality of the operating system of that device o where attempting to download apps from an app distribution service. Age assurance completed for a user account on one ecosystem should ideally be interoperable with other online services to the extent possible, including other ecosystems.

[153] Office of the Australian Information Commissioner, *Privacy Act Review Issues Paper Submission – Part 5: Notice and Consent*, 5.37.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
	For example: • Age assurance completed through one centralised user account could be shared as a user-attribute through a 'single sign on' option on another online service. • Age assurance attributes could be made interoperable across technological ecosystems, to reduce the risk of user lock-in to ecosystems. **Non-ecosystem affiliated online services** Where an online service with a high likelihood that an Australian child end-user will access class 2 material *is not* associated with an ecosystem of products, age-assurance measures could be either: • Implemented by that online service individually in a reasonable and appropriate manner, consistent with existing OSA obligations • Interoperable with centralised user accounts such that age assurance from those accounts can be leveraged for the individual service. **Defaults which may apply if age assurance is not adopted** Where an online service with a high likelihood that a child will access class 2 material on the service cannot or otherwise chooses not to take reasonable steps to confirm end-users' ages, it may, by default, prevent Australian end-user access to class 2 material which is made available on that service. **Interaction with parental controls** Parental controls granting the ability to limit or alter child account access to class 2 materials on services should still be made available as a complement to base age assurance measures and any default settings which may apply as a result of age assurance. **Privacy impact assessment** Industry participants should consider conducting a privacy impact assessment of any measures adopted, to assist with their assessment of both positive and negative privacy impacts of any measures. Privacy impact assessments could be completed with reference to the OAIC's Guide to undertaking privacy impact assessments and Privacy impact assessment tool.
3.0	**Implement appropriate protective measures as a default to child end-users to prevent their access or exposure to class 1C and class 2 material. Apply measures proportionately in line with the suggested approaches to different kinds of material proposed at Table 2 above.**
3.1	RES DIS SMS **Default settings** These services should set default privacy and safety levels to the highest settings available for child end-users to protect and prevent children from being exposed to class 1C and class 2 material.[154]

[154] *Schedule 1 – Social Media Services Online Safety Code (Class 1A and Class 1B Material)*, Measures 6 and 7.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
	Measures and tools to prevent access These services could implement measures and tools including the following: • implementing age-gates on entire services where the primary purpose or function is providing class 2 material; or, on identified areas of services with the primary purpose of providing class 2 material[155] • implementing interstitial notices or functions e.g. warning labels, blurring, halting auto-play, and notice screens on class 2 material which is distributed to child end-users through news and discovery feeds, and where available, through private messaging • filtering class 2 material out of news and discovery feeds by downlisting, deprioritising or quarantining, so that it is not brought to the attention of child end-users • ensuring that recommender systems, algorithms, and other choice architecture, do not promote class 2 material to child end-users • ensuring that end-users are able to report or flag content which they consider may be contrary to a service's terms of use, or is not appropriately tagged as being unsuitable for children, and that these reports are considered and actioned appropriately[156] • ensuring compatibility of services with third-party filtering software or tools which may be installed on devices, or provided by internet carriage services.
3.2	SES **Default settings for safety tools** Search engines should apply safety tools and settings by default (like safe search) for end-users that have not completed an age assurance process, or are otherwise identified as child end-users, to protect and prevent children from exposure to class 1C and class 2 material. **Compatibility with other age assurance and safety tools** Search engines should ensure compatibility with age assurance processes and safety mechanisms associated with other technological ecosystems, to allow end-users to rely on age assurance measures they have already completed to apply to use of search engines. **Reporting** Search engines should provide mechanisms for users to report websites which contain class 1C and class 2 material and do not comply with requirements either under a Phase 2 Code or another regulatory age assurance scheme.[157]

[155] RES, DIS and SMS services which provide online material from Australia are already subject to the RAS Declaration and should already be complying with requirements of this nature.

[156] Effective user reporting systems for dealing with content should be accompanied by the capacity to appeal any decisions based on user-reporting. Reporting appeals should be available both to users who have reported content and believe it has been incorrectly actioned, and for users that may have posted the content which is subject to the reporting and believe it has been incorrectly actioned.

[157] Effective user reporting systems for dealing with non-compliant search results should be accompanied by the capacity to appeal any decisions based on user-reporting. Reporting appeals should be available both to end-users who have reported non-compliant search results and believe they been incorrectly actioned; and also to websites which have been the subject of user reporting and believe that they are compliant with age assurance requirements and that the measures have been incorrectly actioned.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
	Where end-users report that a website containing class 2 material is non-compliant with requirements, search engines should de-prioritise search results of these websites so they are not preferred above compliant websites. **Search advertising** Ensure that search advertising does not serve advertisements for websites which may contain class 1C or class 2 material where a user has safe search activated.
3.3	EQP **On-device measures** On-device safety measures to protect children from access or exposure to class 2 material should be turned on by default. The ability to opt out might be restricted to those who complete age assurance or parental consent processes. **Account profiles** Devices should provide the ability to establish accounts, profiles and age settings for different users (for example, so that an adult user can have a different profile on a device to a child). Child end-user profiles should have the highest safety measures applied as a default. Adult users should be given options to enable safety measures to restrict or limit harmful content for a device which they intend to give to, or share with, a child. **Cost and application** On-device safety and parental controls should be free, and interoperable across all software which can be used on a device.
3.4	HOS **Terms of service** A provider of a third-party hosting service should have and enforce policies and/or contractual terms which apply to customers of the service setting out that they must, when using the service, comply with applicable Australian content laws and regulations to help protect and prevent children from being exposed to class 1C and class 2 material.
3.5	ICS **Easily Accessible User Information** Providers should ensure that Australian end-users are advised of how to help prevent access to class 2 material by child end-users on an ICS, including by regularly notifying them about filter products, including the Family Friendly Filter program. **Safety Tools** Ensure compatibility between internet carriage services provided to end-users and third-party filtering or blocking tools which may be activated by customers of that service to prevent and protect children from being exposed to class 1C and class 2 material.

Matter 1: Protect and prevent children in Australia from accessing or being exposed to class 1C and class 2 material.	
3.6	APP **App review** App distribution services should include consideration of the appropriateness of any developer-submitted age rating as part of any app review process. App distribution services should provide user reporting options for apps which may have been inappropriately age rated. **Enforcement of age ratings** App distribution services should enforce age rating systems by preventing children's access to apps which have been rated as age inappropriate as a default measure. App distribution service search results and advertisements should not serve details of age inappropriate apps to child end-users.

Matter 2: Online industry must provide Australian end-users with effective information, tools and options to limit access and exposure to class 1C and class 2 material.

4.0	**Allow and empower all users to limit their exposure to class 1C and class 2 material.**
4.1	RES DIS SMS EQP SES APP **Safety Tools** Providers should allow all users to opt-in at any time to safety tools which may limit their access or exposure to class 2 material including but not limited to: • filtering material • blocking material • blurring material • halting autoplay of material • placing interstitial notices on material so that users can click through to view if they wish. Providers should ideally allow safety options to attach to a centralised user account which can carry across multiple products and services in a connected technological ecosystem, to provide frictionless access to these safety measures. Safety tools should be made available to end-users which address all different kinds of class 1C and class 2 material. eSafety suggests safety measures should deal with different kinds of material in the terms suggested at **Table 2** above. Services should be compatible with third-party filtering software or tools which may be installed on devices or provided by internet carriage services. **Parental Controls and Options** Services should remove the onus from parents and carers to implement safety features for children by putting safety measures for child end-users in place by default where possible.
5.0	**Provide end-users with control over what they are served through algorithms and recommender systems.**
5.1	RES DIS SMS EQP SES APP **Algorithm and recommender system options** End-users should be empowered to make choices about recommender systems and algorithms which may reduce the occurrence of class 1C and class 2 material in news and content discovery feeds by downlisting, deprioritising or quarantining that material for a user.

Matter 2: Online industry must provide Australian end-users with effective information, tools and options to limit access and exposure to class 1C and class 2 material.	
	User information Information and regular alerts should be provided to end-users about how their personal information may be used to serve them class 1C and class 2 material. This includes information which is: • provided proactively by a user; and • gathered according to the terms of service of an online service.
6.0	**Consistently and measurably improve mechanisms to allow all end-users to better control their exposure to class 1C and class 2 material.**
6.1	RES DIS SMS EQP SES APP **Improvement of protective tools** Services should invest in and measurably improve systems which can detect class 1C and class 2 material and automatically action that material before it is encountered by end-users according to their user preferences. This should include increasing the capability of automated tools to make determinations about material which may be class 1C or class 2, incorporating factors like context. They should measurably invest in and adequately resource: • trust and safety teams dedicated to implementing regulatory requirements and implementing policies which enhance safety for users on online services • moderation teams who conduct human review of flagged material, and can consider material including factors like context while automated consideration of these factors is not yet possible.
6.2	ICS **Improvement of protective tools** Providers should measurably invest in and improve the efficacy and end-user experience with filters and parental controls, to encourage users to adopt these tools and reduce user drop-off from filters as the result of poor service or user experience.

10. Registration process

Code registration

Once a code is finalised by Notice Recipients, an electronic copy must be lodged with eSafety. This should include any submissions supporting the request to register a code or codes, along with any related documentation.

eSafety will use the code and supporting material to assess whether the code or codes should be registered. Supporting documentation should include a range of information, such as:

- the name of the industry association(s) and/or body/bodies lodging the code
- an explanation of the industry section(s) the code is intended to cover
- evidence demonstrating how the industry association(s) and/or body/bodies represent the industry section(s) as defined in the OSA
- evidence demonstrating how the industry association(s) and/or body/bodies have consulted on the code/s with members of the public and industry participants, as required by the OSA.

This could include:

- a list of members of the Notice Recipients who drafted the Codes
- a list of other represented industry participants and how they have been engaged in the development of the code
- a list of industry participants with whom the Notice Recipients have attempted to engage, but who have not provided input into the development of the code
- an explanation as to how the code provides appropriate community safeguards for matters of substantial relevance to the community
- details of the industry and public consultation that was undertaken during the drafting of the code. Information could include a summary of how the consultations were publicised and conducted, a summary of consultation feedback, and how the code has regard to consultation views (e.g., any changes made to the draft code to reflect consultation feedback).

eSafety expects it will require at least four working weeks to review any code(s) lodged for registration.

All registered codes are maintained in an electronic Register that is publicly available on the eSafety website.[158]

Timing

As set out in chapter 2, eSafety has considered industry's preliminary feedback that it requires at least 12 months to produce the Phase 2 Codes. eSafety believes for the reasons set out in that chapter, that its proposed timeframe is reasonable. It reflects the community interest in the timely development of these codes, the work done both in the Phase 1 Codes development process, and the last eight months of industry consultation about the Phase 2 Codes.

eSafety has therefore requested that industry:

- return draft codes as an indicative target of progress by **3 October 2024**
- submit Phase 2 Codes to the Commissioner to consider for registration by **19 December 2024**.

Consultation expectations

Prior to seeking registration of a code, the OSA requires that industry bodies or associations must publish a draft of the code and both invite submissions and consider submissions from:

- members of the public (s 140(1)(e) of the OSA)
- industry participants in the applicable section of the online industry (s 140(1)(f) of the OSA).

Consultation must run for at least 30 days (section 140(3) of the OSA). Consultation with the two groups can be conducted concurrently, as in Phase 1.[159]

General principles

eSafety expects the Notice Recipients to conduct meaningful consultation with both members of the public and industry participants. Consultation should be genuine, widely accessible and transparent. This will help ensure the codes are both capable of being implemented and adequately address community concerns. Meaningful consultation also

[158] eSafety Commissioner, *Register of industry codes and industry standards for online safety*, 26 June 2024.
[159] eSafety Commissioner, *Development of Industry Codes: Position Paper*, September 2021, p. 58.

provides industry participants with an opportunity to engage with, and prepare for, proposed code requirements.

Appropriate forms of consultation could include working groups, focus groups, surveys, web forums and consultation on draft written provisions. A combination of these methods can ensure effective consultation with interested parties. To make sure the consultation process is widely accessible, input should be sought from a variety of relevant stakeholders. Notice of the consultation processes should be published in a way that is likely to reach and encourage submissions from a diverse range of stakeholders. To maintain transparency, eSafety considers that all submissions received as part of the codes development process should be genuinely considered and promptly made available on the industry association's website.

Public consultation

For public consultation, relevant parties in this process may include, but are not limited to:

- children and young people
- children's rights, wellbeing, and protection groups
- parents and carers and their representative groups
- educators and frontline service workers
- domestic and family violence groups
- users of the services and devices (including content creators impacted by the codes) and consumer groups
- groups representing sex workers
- digital rights and civil society groups
- representatives of the LGBTIQ+ community
- representatives of culturally and linguistically diverse communities
- First Nations community representatives
- representatives from academia
- the safety tech sector.

Some stakeholders may participate in both public and industry consultations. eSafety encourages industry associations to consider the voices and experiences of diverse and at-risk groups in developing industry codes, including those with lived experience of the harms the codes seek to address.

Invitations to provide input on a draft code should provide easy and accessible access to the draft, be likely to reach the intended groups, use plain language to explain the code's purpose, and clarify its key issues. The invitation should also explain how the public can contribute to the development of Phase 2 Codes.

Bearing in mind the six-month timeframe for developing the Phase 2 Codes, eSafety also encourages industry to extend the consultation period beyond the minimum required in the OSA if possible.

Industry consultation and representation

Industry associations must take steps to ensure the Commissioner can be satisfied they have completed consultation requirements set out in sections 140(e)-(g) of the OSA. In assessing this as part of the code registration process, eSafety will take into consideration the scale and breadth of the industry consultation conducted. This is crucial for developing the Phase 2 Codes, as they will affect substantial sections of the online industry, such as major pornography sites, which are not currently within the membership of Notice Recipients.

When submitting codes to the Commissioner, submissions should include details about the nature and scope of the consultation with relevant industry participants, such as providers of mainstream online pornography (discussed above), as well as local Australian providers of online pornography. These submissions should demonstrate the adequacy of the consultation process to ensure a representative industry code. Information to be included with these submissions to eSafety should detail the number of meetings or consultations held with relevant industry participants who are not currently members of the Notice Recipients.

eSafety also encourages online pornography providers to proactively engage in this process. This presents an opportunity for these providers to offer valuable feedback on measures they believe may best address the matters to be dealt with. eSafety expects that these providers can also provide the Notice Recipients with valuable insights to help inform the best measures to adopt in the Phase 2 Codes.

It is important for online service providers distributing class 1C and class 2 material to end-users in Australia to understand they *will* be subject to the requirements of any registered industry codes, regardless of whether they have actively participated in drafting those codes. For example, if submissions on consultation show that Notice Recipients have made reasonable efforts to secure the views of relevant industry participants who have not meaningfully engaged with their processes, the Commissioner will take this into account in her decision to register the Codes.

The content of these submissions and the arguments presented will influence the Commissioner's assessment of whether the Codes provide appropriate community safeguards in relation to the two specified matters. These considerations are essential for the Commissioner to be satisfied to register the Codes.

Consultation with eSafety

Section 140(1)(g) of the OSA requires the Commissioner be consulted about the development of industry codes.

Most industry groups in the steering group that drafted the Phase 1 Codes have engaged with eSafety about the Phase 2 Codes except BSA, which has, to eSafety's knowledge, withdrawn from the Phase 2 code drafting process as set out in chapter 5.

eSafety acknowledges industry's early engagement in informal stages of Phase 2 Codes development and looks forward to future efforts in response to the formal section 141 Notices issued for Phase 2.

eSafety expects industry associations, and individual companies, to continue to engage with eSafety throughout the codes process. This engagement will help track code progress, and address any areas of overlap, duplication, or other concerns promptly. It will aid future code implementation, promote consistency and coordination across codes, and ensure the registration stage runs efficiently and effectively. Early engagement also helps eSafety identify potential gaps in community safeguards, which may necessitate the need to develop Industry Standards.

Administering and reviewing the Codes

eSafety believes industry should adopt consistent principles for administering and reviewing the codes, akin to those adopted in Phase 1.

Next steps

As discussed in chapter 5, the section 141 Notices request that industry return proposed draft Phase 2 codes, accompanied by submissions and supporting material outlined above, to eSafety by COB Thursday 19 December 2024.

eSafety will continue to engage with industry throughout the code development process. The industry bodies drafting the Phase 2 Codes will also conduct public consultations as the drafting proceeds.

eSafety thanks industry in advance for its efforts to prepare the drafts of these important codes over the next six months.

Annexure A – case studies

Case study 1 - Jeffrey

Jeffrey receives a new smartphone for his 13th birthday. During set up, he is prompted to create an account. This requires him to enter his birthdate, which he does accurately, and provide credit card details, which he gets from his parents. The account is necessary to download apps and enable the full functionality of the phone, including its opt-in parental controls and safety and privacy settings. Jeffrey's parents are not tech-savvy, and are unaware of these features, so the phone remains without controls enabled.

Jeffrey connects to the internet using his home WiFi, which also lacks any controls. Using the default browser on his phone, he enters a search term recommended by a friend. Although the search results page offers an option to enable safe search, Jeffrey ignores it. Among the top results, he clicks on a link to a popular pornography website.

The website's terms of use state that users must be 18, but there are no measures in place to enforce this. On the home page of the website, Jeffrey encounters a large volume of graphic videos, including those depicting and promoting sexual violence. He has never seen this kind of content before and has no information from his parents or school about pornography. He also has a limited understanding of sex, and what respectful relationships look like.

Measures that could be applied to protect and support Jeffrey and other children include:

- **Providing effective and easy-to-apply controls for parents, carers and any other end-users who want to avoid this content**, such as:
    - **Equipment controls:** These may be built into the operating system (e.g., on-device settings to automatically filter web content or limit access to adult content if enabled) or installed separately.
    - **Account level controls:** Filtering and safety settings may be available to users logged into a centralised user account for accessing various interconnected products or services. These settings could be automatically applied when the user accesses products or services across the ecosystem, such as app distribution and messaging services.
    - **Controls provided or supported by ISPs**: These may be integrated into networks, modems, or routers, as well as third-party filters such as those certified under the Family Friendly Filter scheme.

  - **Search engine controls:** This could include safe search functionality to filter or blur sexually explicit content.

- **Providing accessible online safety information for Jeffrey's parents** to make them aware of these controls.
- **Enabling some controls by default** for someone Jeffrey's age, so they are applied regardless of his parents' awareness.
  - **There could also be restrictions on opting out**, such as measures to confirm a user is old enough to opt out of default settings or measures to obtain parental consent.
- **Implementing appropriate measures at the website-level** to support these controls, including for example:
  - **Ensuring compatibility between the website's coding and parental control and filtering tools,** which may be employed to protect children at the device-, account-, or other levels.
  - **Providing a pornography-free home page** that clearly signals users must be 18+ to enter the site.
  - **Taking reasonable and appropriate steps to confirm users are 18+** before allowing them to access pornography (consistent with other expectations under the Online Safety Act). This could include accepting age checks that have already been conducted at other levels (e.g. device-level).
- **Enforcing website-based measures**, such as:
  - **Search engines could de-prioritise websites that refuse to apply measures** in search results ranking them lower than compliant websites.
  - **Hosting services could enforce compliance with terms of** service and cease to host websites that refuse to comply.
- **Providing information to help Jeffrey understand online pornography** and supporting his parents and educators to speak to him about this issue.

Case study 2 – Sally

Sally (aged 14) uses a hand-me-down smartphone which her parents gave to her when they upgraded their own devices. During the phone set up, Sally's parents linked her account to their own, using her real birthdate. They enabled some controls for web browsing, but not for app downloads.

Sally has downloaded several social media, gaming, and messaging apps. Some of these apps have age ratings of 13+ and others have age ratings of 17+ due to the nature of the content they permit. However, these age ratings are not enforced during download.

When Sally signs up for accounts on these services, she provides a false birthdate indicating she is 18. None of the services take steps to confirm her age during sign-up.

As Sally scrolls through her feed on one of the services, she is served with content from an influencer. She decides to follow the account, which posts a variety of high-impact material involving sex, nudity, violence, drug use, and suicide. Some of the content violates the service's rules and is occasionally removed, while other posts are allowed but should be tagged as sensitive and shielded from younger users. Sally begins to spend more time viewing the influencer's content and starts receiving recommendations for increasingly extreme content through her feed. She also receives direct messages from accounts offering access to explicit content on other services. This makes her feel upset and uncomfortable, but she is not sure what to do about it.

Measures that could be applied to protect and support Sally and other children include:

- **Accessible online safety information:** Provide resources to make parents and carers aware of the types of material and activity children may encounter on various apps, and how to support them through open conversations and technological controls.
- **Enforcement of app age ratings by default:** Ensure children cannot download apps that contain content that is inappropriate for them.
    - **There could also be restrictions on opting out**, such as measures to confirm a user is old enough to opt out of default settings or measures to obtain parental consent.
- **Taking reasonable and appropriate steps to confirm users' age** before allowing them to access high-impact material (consistent with other expectations under the Online Safety Act). This could include accepting age checks that have already been conducted at other levels (e.g., by the app distribution service), preventing the need for multiple age checks and minimising data collection.
- **Establishment and enforcement of rules on services**, for example:

  - Where certain types of content are not permitted – and there are ways to **detect such content proactively**, such as nudity classifiers – services should apply these options. There should also be **accessible and easy ways to report** content that is not permitted. Both methods should be supported by **human review**.
  - Where content that is not appropriate for children is permitted – and children are also allowed to use the service – rules need to be applied and enforced to prevent children from encountering inappropriate content. This includes **marking harmful material and restricting its visibility and access to children**.
  - Services should set and **enforce minimum age requirements**, including by taking steps to detect (including through reporting) end-users whose accounts have likely been set to a false age, and restricting or removing access to accounts until their age is confirmed.

- **Providing age-appropriate experiences**. This includes enabling robust safety and privacy protections by default for children, including restrictions on who can contact them.
- **Providing user empowerment tools.** This could include tools to blur sensitive content, place it behind interstitial notices, and allow end-users to control content in their feeds or searches. End-users should also be able to restrict who can message them.

For more information on the specific measures eSafety suggests could be adopted in the Codes across the eight sections of the online industry, see chapters 8 and 9.